655915

COMPUTER-ASSISTED EXPLANATION

Edmund C. Berkeley

DDC
RECEIVED
AUG 9 1967
C

Distribution of this document is unlimited.

Information International, Inc.
Cambridge, Mass. 02139

RECEIVED
AUG 11 1967

"COMPUTER-ASSISTED EXPLANATION".

First Printing

ERRATA as of May 12, 1967

p. v, 3rd paragraph, line 2: Replace "photoffset" by "photo-offset"

p. 3, lines 6 and 7: Underline "stern"; delete underline on "bow"; underline "forward".

p. 6, line 4: Insert "?" after "Explaining"

p. 10, 3rd paragraph, line 3: Replace "hard to" by "hard for"

p. 17, line 17: Replace "scientist"s" by "scientist's"

p. 43, line 6 from bottom: Insert quotes (") after "or"

p. 64, line 9 from bottom: Replace "and as" by "and so"

p. 65, line 6: Delete comma after "reading"

p. 69, line 2: Delete comma after "nad"

p. 69, line 15: Replace "three 3's" by "four 3's"

p. 72, line 6 from bottom: After "what is" insert "in"

p. 73, line 6: Replace "made" by "make"

p. 73, line 7: Replace period (.) by colon (:)

p. 89, line 14: Replace "focusses" by "focuses"

p. 151, line 9 from bottom: Insert quotes (") at end of line

p. 262, line 4: Insert quotes (") after "composition"

p. 265, line 4 from bottom: Replace "Bolean" by "Boolean"

p. 269, at "Basic English" entry: Replace "57" by "35, 57"

p. 270, at "Ogden, C. K." entry: Replace "57" by "35, 57"

COMPUTER-ASSISTED EXPLANATION:

A Guide to Explaining: and Some Ways of Using a Computer to Assist in Clear Explanation

EDMUND C. BERKELEY
Senior Scientist
Information International, Inc.

Distribution of this
document is unlimited.

May, 1967

Information International, Inc.
545 Technology Square
Cambridge, Mass. 02139

Acknowledgments

The work reported here was made possible by Contract N00014-66-C0281 from the Office of Naval Research to Information International Inc.

Grateful acknowledgment is made to the following persons who contributed much in the form of illumination, ideas, suggestions, and content: Edward Fredkin, Stewart Nelson, Daniel J. Edwards, Glenn L. Bryan, and John A. Nagay.

Grateful acknowledgment should also be made to the persons who prepared the final photoffset master copy for this book — Barbara Chaffee and Evelyn Frank.

Edmund C. Berkeley

Introduction

The Subject and Purpose of This Book

The Subject of This Book

The subject of this book is explanation. To explain something is to make it clear and plain, to change it from something that is not understood to something that is understood.

Explaining involves something to be explained (a topic), a person who does the explaining (the explainer or lecturer), persons to whom the explanation is given (the audience), and many more elements and factors besides. To explain well is an art, and the capacity to explain well can be learned.

Assistance from a Computer

The advent of the automatic computer makes it possible to apply to the art of explaining a vast power to handle information. The programmed computer, for example, can easily provide a great deal of control over many factors such as vocabulary, sentence length, and number of syllables. The result is that better explanations can be produced with less human labor. In addition, in the next few years, the programmed computer should become able to deal efficiently not only with words but also with ideas. The computer should be able to deal intelligently with terms and relations, statements and questions, and thus provide explanation ideally suited to each individual member of an audience. Many other interesting and important possibilities for computer-assisted explanation clearly cast their shadows before.

The Purposes of this Book

This book has two purposes. The first is to examine what is good explaining, and to present a guide to the art of explaining. At present, in most cases and for most tasks, in producing good explanation, we cannot yet use the resources of the automatic computer, and human beings have to produce the explanation. The second purpose, however, is to demonstrate a beginning in using an automatic computer to help produce good explanation. Five working computer programs are stated in this book.

The Contents of this Book

What is explaining, and what isn't? What is good explanation, and what isn't? The first few chapters of this book discuss these questions.

The quality of an explanation is affected by a large number of factors. What these are is discussed in Chapter 4.

When we are setting out to explain something, we can usefully recognize four classes of words. One of these classes, the Special Vocabulary, has long been recognized. I call the other three classes the Allusion Vocabulary, the One-Syllable Vocabulary, and the Key Vocabulary. How to make these distinctions, and apply them usefully in producing good explanation is discussed, mainly in Chapters 5 and 6.

Some of the possible applications of programmed computers to explaining well are described, and programs are given for them, in Chapters 6, 10, and Appendix 1.

Ways of improving some actual sample explanations are shown in Chapters 7 and 8.

The value of improvement in explanation and some other practical questions are discussed in Chapter 9.

The last chapter is a short summary of advice for explaining well.

The Role of the Human Being

At the present stage of the art of programming computers, however, a programmed computer by itself cannot compose clear and satisfactory explanation of any given subject for human beings. At best, it can only help human beings compose explanation in English.

Particularly when revising an explanation to improve it, we have to use human beings to compose revision, and not a computer. This role for human beings may continue for some time.

The Future

Explanation, it seems, is a new field for study and for applying the scientific method; and it seems to be one in which much significant progress can be achieved if we judge by the level of explanations commonly seen nowadays. Consider for example the many times when, as we read something, we find that we fail to understand what is being said, and are perplexed or blocked.

The present report merely touches the surface of this field of explaining well; a great deal of cultivating is needed in the future.

Any comments, criticisms or suggestions that any reader cares to send me will be gratefully received.

Edmund C. Berkeley
Information International Inc.
545 Technology Sq.
Cambridge, Mass. 02139

May, 1967

CONTENTS

Acknowledgments		v
Introduction		vi
Chapter		
1	For Instance?	1
2	Explaining	6
3	The Differences Between Explanation and Related Subjects	13
4	Factors Affecting the Quality of an Explanation	22
5	The Classes of Words for Explaining	54
6	Vocabulary Analysis with a Computer	66
7	Judging and Improving an Explanation	75
8	Judging and Improving Two More Sample Explanations	127
9	Practical Aspects of Revising and Improving Explanations	154
10	Other Assistance from a Computer in Explaining	159
11	A Short Summary of Advice for Explaining	179
Appendix		
1	Computer Program for Vocabulary Classification using a DEC PDP-7 Computer	181
2	Eight Sample Explanations and their Assessment	205
3	"Giving Power to Words" by Philip W. Swain	252
4	Bibliography	260
5	A Short Glossary	266
6	Index	269

Reproduction in whole or in part

is permitted for any purpose of the United States Government

Copyright © 1967 by Information International Inc.

This research was sponsored by the Personnel and Training Branch, Office of Naval Research, Department of the Navy, Washington, D. C., under Contract N00014-66-C0281.

First printing, May, 1967

Printed in the United States of America

Chapter 1

For Instance?

Before we launch into an extended discussion of the subject of "explanation" and its applications and implications, would it not be good to give some examples? examples of good explanation and poor explanation, so that you the reader may see some actual instances of what we are discussing in this report?

1. Notice in a Taxicab

Following is a notice that I read in April 1949 in a New York taxicab:

AVOID INJURY
ACCIDENTS

SIT BACK -- RELAX -- ENJOY YOUR RIDE

This is excellent explanation. The words are good, direct words, with clear meaning to almost everybody who rides in a taxicab. The entire explanation is short: 9 words in all. The reason for the directions is given first: "avoid injury and accidents". Then the directions are stated, "sit back and relax". Then one more encouraging reason is given, "enjoy your ride."

2. A Paragraph from a Scientific Article

The following is part of the third paragraph from an article "The Significance of Meiosis in Allomyces" by Ralph Emerson and Charles M. Wilson in Science, July 22, 1949, vol. 110, pp 86-88:

Resistant sporangia formed by sporophytic thalli grown on slants of yeast starch agar ordinarily become capable of germination three to six weeks after their formation. At this time each sporangium contains about a dozen expanded, diploid nuclei in an advanced prophase stage. These sporangia are fully mature and, if air dried, they will remain viable and their nuclei will remain in prophase without any further detectable change for periods up to at least ten years. When mature resistant sporangia are taken directly from moist agar cultures and placed in water at 20° to 25° C., they form and release spores in 100 to 130 minutes. During this short interval the two meiotic nuclear divisions occur, and are immediately followed by cleavage of the cytoplasm and organization of the zoospores. Each of the zoospores is haploid and normally uninucleate. ....

This is straightforward technical writing of the ordinary technical journal variety. Also, the writers have pointed out something remarkable -- here is a form of life which has a resting period ten years long, and then in two hours' time under the right conditions it lives actively again and makes spores. But the following words, the special terms, all have to be understood in order that the paragraph have full meaning:

agar	germination	sporangia
centigrade (i.e., C.)	haploid	sporangium
cleavage	meiotic	spores
cultures	nuclear	sporophytic
cytoplasm	nuclei	thalli
diploid	prophase	uninucleate
divisions	resistant	yeast starch
expanded	slants	zoospores

fact, out of 99 different words in this paragraph, 24 are special terms. This is undoubtedly much too heavy a load for at least some readers and probably a great many.

3. A Paragraph from a Navy Guide for Seamen

Following are the first two paragraphs from page 210, Chapter 12, "Facts about Ships", in "The Bluejackets' Manual", 17th edition, published by the U.S. Naval Inst., Annapolis, Md., 1963:

> A man does not get on a ship; he goes aboard. The front of a ship is the bow, the rear end is the stern. When a man stands at the center of the ship and faces the bow, he faces forward; if he turns around, he faces aft. Facing forward, the right side of the ship is the starboard side; the left side is the port side. An imaginary line from bow to stern is the centerline; it runs fore-and-aft. The length of this line is the length of the ship. The greatest width of the ship is the beam.
>
> An object directly off the side of the ship is abeam. An object or line running directly across the ship, like a passageway or deck beam, is athwartships. A man standing at the center of the ship is amidships. When he faces either side, he faces outboard. His shipmate at the rail who looks back at him is facing inboard. An object over his head is above; an object underneath him is below.

These two paragraphs contain 86 different words, of which 18 are being defined in the course of these two paragraphs. Yet what a great contrast with the last explanation -- what a great amount of help is given to the person interested in understanding! There is only one mistake in explanation: the word "passageway" is assumed to be familiar to the audience but is not.

Furthermore, these two paragraphs as a whole are entertaining -- because in a light vein they make clear how different language aboard ship is from language on shore. This kind of explanation can be read for enjoyment as well as instruction.

4. Two Paragraphs from a Navy Publication for Officers

The following two paragraphs are taken from pages 40 and 41 of Chapter 8, "The Case of the Suspended Sentence", in "Principles and Problems of Naval Leadership", published by the Bureau of Naval Personnel, 1964. The audience is assumed to be Navy officers who are interested in leadership and the supervision and management of men. Following are the first two paragraphs of discussion presented in answer to the following question:

Why doesn't the Navy screen prospective recruits to keep the troublemakers out of the Navy?

This question appears earlier in Chapter 8.

SCREENING OUT POTENTIAL TROUBLEMAKERS

If the behavior of human beings were wholly determined by heredity and environment, or by purely mechanical or automatic laws, prediction of future success or failure would be easy. If we could say that human beings when stimulated in a particular fashion, would automatically react in a given way, not only the predictability but the control of human behavior would be easy. Indeed some schools of thought and teaching have given this impression. One such school might loosely be termed "behavioristic", another "mechanistic". Developments in the science of cybernetics (a comparative study of complex mechanical-electronic communications systems, such as computers, with the control system formed by the nervous system and the brain) have led some people to believe that there is no essential difference between men and highly refined "computers".

But the fact of the matter is that no machine ever invented past, present, or future, no matter how complex its operations, has had or can have a built-in personal moral responsibility. A man can know all the rules. He can study behavioral principles all his life. In turn he can be

studied in exhausting detail. His blood can be tested; x-rays and fluoroscopes can strip him of every physiological secret. Electro-encephalograms, psychogalvanometers, and Rorschach diagnoses can tell a lot about what goes on in his mind and in his emotional make-up. Chi-squares and coefficients of correlation can ascertain the probabilities of how many people will prefer blue wallpaper to red wallpaper.

In this case, we have now read two paragraphs and have not yet found any direct connection with the subject, "Screening out Potential Troublemakers". Furthermore, the audience probably will not understand behavioristic, electroencephalograms, psychogalvanometers, Rorschach diagnoses, chi-squares, and coefficients of correlation. Also, many of the statements are clearly not true: blood tests, x-rays, and fluoroscopes are not enough to find out every physiological secret of a man. Many of the terms are unclear: what is a "highly refined computer"? Although the vocabulary is not as a whole difficult for this audience, we cannot say that this passage is good explanation.

5. The Main Argument

These examples of explanation illustrate what will be the main argument of this report:

1) Much writing has the goal of explaining something to an audience.

2) Much explanation is poor.

3) It should be possible to improve greatly the quality of explanation.

4) And it is possible.

Chapter 2

Explaining

Outline

1. What is Explaining
2. The Elements of Explanation
3. The Result of Good Explanation
4. The Importance of Explanation
5. The Factor of Advantage to be Gained from Good Explanation
6. Advances in the Techniques of Explanation
7. Computer-Assisted Explanation

Everybody has the task from time to time of explaining something to somebody else. It is a common problem, from telling a stranger which way he should go, to telling an employer why you are the right person to be employed for a certain job, to instructing a new member of the United States Navy in the kind of work he should do or the nature of concepts he should learn. How many of us can explain well? How many of us are aware of all the factors that may affect the success of an explanation?

This report seeks to make clear many of the factors that affect successful explanation, and a number of other considerations that bear on it. It also seeks to make clear a number of the ways in which automatic computers now or in the future may assist in producing successful explanation. If you become aware of the factors that may affect successful explanation, you are likely to be able to explain better, for knowing that a factor exists is the first step in acquiring control over it.

1. What is Explaining?

To explain is to make plain or clear, to render intelligible. To explain a strange idea means to define that idea in terms of other, more familiar ideas; it means saying exactly what the strange idea is, using familiar ideas and familiar relations.

For example, what is fatigue duty? One explanation (taken from the dictionary) is that "fatigue duty" is a military term, and it is:

> occasional work performed by selected details of soldiers, in addition to drill duties, especially policing, painting, and camp maintenance.

This explanation tells the context of "fatigue duty". It then uses some familiar ideas such as "work, soldiers, painting", and some less familiar ideas such as "details of soldiers", and "drill duties". These if not known can be fairly easily guessed. Then the whole set of ideas is put together in the familiar relation:

> ".... is .... performed by ....., in addition to ....., especially ....., ..., and ... ."

This is good explanation for an audience of a great many people, particularly the kind of people who might want to know what "fatigue duty" means; and this kind of explanation marks a good dictionary.

The opposite process, expressing familiar ideas in terms of strange ones, is illustr ed by a famous definition attributed to Dr. Samuel Johnson:

> "Man is the featherless biped."

In this case a very familiar idea is defined in terms of two much less familiar ideas. Clearly, this is far from being good explanation for anybody; we can conclude that the definition is not intended seriously.

2. The Elements of Explanation: Topic -- Explainer -- Audience -- Vocabulary -- Presentation

In explanation we can usefully distinguish five elements. The first is the thing

which is being explained; let us call it the topic or subject of the explanation. Second is the person who is presenting the explanation; let us call him the explainer or the lecturer. Third is the person or persons who are receiving the explanation; let us call them the audience. If there is only one person, we may call him the explainee. Fourth is the collection of words expressing familiar ideas and relations by means of which the explanation is presented; let us call this the vocabulary. The fifth element we may call the presentation, which consists of the vocabulary organized by the lecturer to explain the subject to the audience.

In general we shall here consider the explaining of technical and scientific subjects. The lecturer, we shall assume, is very well informed about the subject, and somewhat interested in trying to explain the subject as well as he can, provided that that effort does not take too much time away from his main interests, which are the subject and not the explanation. The audience we shall assume consists of persons who have some interest or curiosity in the subject, but they have other matters also that they are interested in; in fact, we can imagine them looking at their wrist watches and not eager to stay any longer than necessary or desirable from their own point of view.

Whenever a piece of exposition is prepared for explaining some subject to an audience, the explainer has to make decisions about the knowledge that his audience has: what they know, and what they don't know.

A speaker has a great advantage. He can say to his audience, "If any of you do not understand some term that I am using, please speak up and make me explain it."

A writer does not have this advantage. Accordingly, he needs to judge what his audience knows and does not know, and take that judgment into account when he is writing. To carry this out well, a writer has to notice every word and phrase that he uses, and ask himself for each one:

Will my audience understand this?

It is not hard to do this checking for a paragraph or two, but for a chapter or a

manual or a book, it is a great deal harder, and requires a lot of effort and detail. Here is one of the places where a computer can help.

In general, an explainer should watch every word and phrase that he uses for explaining. He should keep track of these terms. And whenever he finds that he needs to use a term which his audience perhaps does not understand, then he must give more information about the term, so that the audience can surely understand it.

A computer can be a substantial help in this process. It can compare the words and phrases that a writer is using and the vocabulary that the audience is assumed to understand; then it can report those words and phrases which occur in the writing but which are not in the audience's vocabulary.

3. The Result of Good Explanation

When an explanation is successful, we fully understand. What is understanding?

Most of the definitions given in the dictionary for "understanding" do not convey an operational meaning; they are synonyms like "grasping", "comprehending", etc. But among the definitions there is one based on operations that can be observed and tested: understanding is "the power to distinguish truth from falsehood and to adapt means to ends".

Accordingly, if an explanation about some idea is good, is effective, we understand. As a result we can distinguish true statements about that idea from false ones, and we can behave in regard to that idea in such a way as to adapt means to ends. For example, if we understand directions given to us about how to go to some destination, we prove our understanding by arriving there.

The process of understanding is general. In other words, it is just the same process for every new idea to be understood. We find out the name of the idea. We collect true statements about it. We practice applying the new statements. The more true statements we have gathered, and the more practice we have applying them, the more we understand the idea.

It is clear that this process can be applied to any idea (that makes sense), and therefore we can understand any idea (that makes sense). Good explanation takes us there rapidly, and poor explanation takes us there very slowly if at all.

4. The Importance of Explanation

As nearly everybody nowadays realizes, we are living in a time when a great tide of new technology is pouring forth. For example, it has been estimated that over 1.0 million jobs in the United States disappear each year as a result of automation and other technological changes.

The new technology and the changes it is producing are placing a heavy requirement on human beings to understand new information. In fact, so many new developments are happening so fast in so many places, that it is very hard to most people to keep up.

In the United States Navy, the importance of explaining, conveying understanding, through systematic education, instruction, and training, has long been recognized. Not only in the areas of systematic education but also in the areas of nonsystematic education, however, it is important to convey understanding from one human mind to another, to make use of good explanations.

5. Factor of Advantage to be Gained from Good Explanation

Because of the fact that not much attention has been paid to techniques of explanation in recent years, the application of the scientific method in this area might well yield a large gain, perhaps by a factor of 2 or more. In other words, good explanations achieved through the application of modern scientific methods might shorten the time needed to gain understanding of something unfamiliar, by 50% or more.

6. Advances in the Techniques of Explanation

Considerable advances in the techniques of explanation are possible. They may be divided into two categories, those that involve human beings only, and those that

involve human beings together with automatic computers.

In the first category, more than forty factors affecting the quality of an explanation may be inventoried. Examples of some of these factors are:

-- Ideas assumed known by the audience: extent to which they are carefully and consistently selected and identified

-- Common properties and relations: extent to which they are reported for a new idea being explained

-- Examples: way in which an explanation uses examples to the optimum extent

-- Words with ambiguous or obscure meanings: absent as far as possible

-- Proportion of short words (words of one syllable) used for explaining relative to long words (words of more than one syllable): optimum proportion

These factors are discussed at length in Chapter 4.

7. Computer-Assisted Explanation

Automatic computers however now exist, and they can be applied to the problems of producing explanation.

A computer can help the explainer, by applying tests and evaluations to his drafts of an explanation; it can signal ways in which the drafts can be improved. With this aid, an explainer can more easily produce good explanations. Research using a computer in this way has been carried out and is reported and discussed in Chapters 6 and 10 of this report.

A computer (especially when it is "time-shared") can also help an explainee by responding to his questions, and behaving like his personal lecturer or explainer. Within a certain subject area, it may have almost unlimited power to remember, calculate, be logical, and evaluate how the explainee is responding (though it may be difficult to program it to respond to the meaning of freely worded statements by the

explainee). This kind of computer use closely resembles what is called "computer-assisted instruction" or CAI; we do not try in this report to cover that subject also.

It is clear from the evidence in this report and the results and direction of the research here reported that the computer provides an additional power for attaining successful explanation: computer-assisted explanation.

Chapter 3

The Differences between Explanation and Related Subjects

Outline

1. Explanation as a Subject (or Discipline)
2. The Related Subject: Instruction
3. The Related Subject: Exposition
4. The Related Subject: Technical Writing
5. The Related Subject: The Presentation of Mathematical Systems
6. The Distinguishing Characteristics of Explanation

1. Explanation as a Subject (or Discipline)

Explanation has not been recognized as a branch of knowledge, or a territory of science or a technical art or a discipline, even though closely related subjects have been so recognized. Among these closely related subjects are:

1) public relations (there are courses in public relations)
2) advertising and propaganda (there are courses in advertising if not propaganda)
3) journalism (there are schools of journalism)
4) homiletics (the art of writing sermons)
5) exegetics (the science of interpreting the Scriptures)
6) education (there are schools of education)
7) writing (there are courses in many kinds of writing)
8) instruction and teaching (there are teachers colleges)

9) exposition (a subdivision of English composition)

10) technical writing (there is now a Society of Technical Writers and Publishers)

11) the presentation of mathematical systems (such as Euclid's Elements of Geometry, which has long been an influential model for explanation)

The first five subjects listed above are clearly different from explanation because they cover a limited territory, whereas explanation relates to any territory at all. The next two subjects, education and writing, are very broad and large subjects:

-- the education of human beings can take place in many, many ways;

-- writing includes explanation and a great deal more besides.

But for the last four subjects, it is worth making quite clear the similarities and the differences. In this way we can indicate why explanation is worth attention as a subject in itself.

2. The Related Subject: Instruction

Instruction and explanation grade smoothly into each other in some ways but in many ways they are quite different.

Instruction regularly implies that there is a deliberately arranged situation in which some persons are students, and another person is the teacher, and explanation is taking place in a deliberately educational environment such as a classroom, a seminar, a lecture, etc.

In this environment there exist several kinds of compulsion. The student is under pressure of "having to go to school" or "being in school". If he is motivated, he wants the end goal of passing courses and graduating with a degree. He knows he must study his homework before the next meeting of the group, which is often tomorrow;

he cannot procrastinate more than an hour or two. He does not want to be laughed at by his peers at the next meeting if he should give stupid answers to the teacher's questions. And so on.

Another kind of instruction occurs when a student is working by himself. He may be in a laboratory session of a scientific course; he may be going through a series of steps in a programmed instruction book; he may be interacting with a computer in one of the varieties of "computer-assisted instruction"; he may be reading a textbook and trying to determine for himself the significant things that are being taught.

In solitary instruction also there is usually strong motivation or compulsion, and in addition there are supplementing resources. He can look up what is not clear to him in some other text. He can search for a teacher who can explain to him what he does not understand. He can discuss questions with a fellow student so that both of them can come closer to comprehending what is being said. Basically, a student is under a pressure to try to learn.

The instructional environment is however missing on many occasions when explanation is needed. For example, an author may be writing a popular exposition of a scientific subject, and he has to rely on the attractiveness with which he may be able to present his subject, in order to persuade his readers to continue reading what he has written. Or a subordinate may be presenting some proposals to a busy supervisor, and he has to rely on the interest and convincingness of his first few explanatory statements to persuade his supervisor to listen to the rest of what he has to say. Or the manager of a project may have to explain to a committee that hold the purse strings why the project continues to be important, how this year's results have been significant, and why a larger budget than last year is reasonable; and he knows he has to rely on some ten pages of explanation which the committee will read in his absence. Or a department of the Government is interested in increasing the skills and understanding of men working for it, and desires to make instructional material, manuals, etc., as clear, interesting, and inviting as possible so as to persuade these men to use some

of their spare time to learn more and thus become more useful. In each of these cases, the audience for the explanation is under no compulsion to listen.

As a result, explanation must be more interesting, more attractive and more "explanatory" than much instruction. But many of the principles of good instruction apply to explanation. A person who has much explanation to write should make himself competent in the technique of instruction.

3. The Related Subject: Exposition

Exposition refers to the principles for informing which are stated in courses in English composition taught to students in school and college. These principles deal with outlines, sequences of points, unity, coherence, emphasis, etc. Perrin's "Writer's Guide" (see the bibliography) states under the heading "expository writing":

> Writing that is intended primarily to inform its readers -- by presenting facts, giving directions, recording events, interpreting facts, developing opinions -- is expository. People do not sit down to write "exposition"; the word is convenient for grouping various articles that convey information.

From this definition it can be seen that explanation is much more specific than exposition. But all the principles of good exposition apply to good explanation. A person who has much explanation to write should make himself a master of exposition.

4. The Related Subject: Technical Writing

Technical writing is often thought of as the kind of writing which is produced to accompany a piece of equipment or a product so that users can operate it and maintenance men can service or repair it. Technical writing also includes writing by scientists, engineers, or professional men which communicates technical information to their colleagues, persons who are assumed to know almost as much about what is being communicated as the author himself.

Technical writing has been defined as follows:

> "Writing in any form is a means of communication between a writer who possesses certain information and a reader who might be interested in that information. Technical writing involves material of a specialized nature, meant for a specific reading group. When the material is properly written, it can be readily understood and utilized by those readers.
>
> -- From "Technical Writing Techniques" by Joseph Racker, Prentice Hall Inc., Englewood Cliffs, N.J., 1960, p. 1

Then the author gives as examples: (1) instructions for assembling a model aircraft from a kit, which would enable the purchaser of the kit to assemble it; (2) a book on the repair and adjustment of hydramatic systems, which would enable average automobile mechanics familiar with standard gear shift systems to repair and adjust hydramatic systems with a minimum of special tools; (3) an engineer's paper on an improved triode tube, which would enable other design engineers to decide whether the new triode would meet their requirements; (4) a scientist"s paper on his new theory and experiments on radio interference by solar and cosmic sources, which would enable other scientists to understand the new theory and to judge how it is substantiated by experimental data.

From these definitions it can be seen that technical writing is essentially a subdivision of explanation in an applied context, usually in a situation where:

(1) there is a captive audience, which is forced by circumstances to do its best to understand; or

(2) there is a very high-level audience which is not required to understand thoroughly, but only to make reasonably good judgments based on incomplete information and approximate understanding.

Thus technical writing regularly does not include extensive explanation of the basic structure or fundamental principles of a subject; and so the information presented in technical writing is often inaccessible to the uninitiated in that area of knowledge.

Even so, a person who has much explanation to write should have a knowledge of technical writing.

5. The Related Subject: The Presentation of Mathematical Systems

The presentation of a mathematical system regularly uses certain simple ideas and axioms at the beginning, and proceeds step by step to many more definitions, and a large number of propositions.

In geometry, for example, the usual systematic presentation recognizes and explains:

Concepts and Statements	Examples
simple ideas (basic concepts -- undefined terms)	point, line, distance, angle
axioms (assumptions, postulates -- unproved statements)	Two lines intersect at a point.
later terms (defined terms)	circle, rectangle, length, width, area
later propositions (proved statements)	The area of a rectangle is equal to the length times the width.

In this kind of mathematical presentation, each later idea is carefully defined in terms of earlier ones, and each later proposition is carefully proved by logical argument from earlier ones.

This kind of mathematical presentation is a mathematical ideal of clear, economical, and beautiful explanation. The Greek mathematician Euclid in his "Elements" did this for geometry. For more than 1500 years, and for more than 1000 editions of his work, Euclid's style of explanation has been a model for many kinds of explanation.

Many topics can be explained well using the methods of presentation of mathematical systems. In fact, it is almost inevitable that a good explanation has to start somewhere, with basic ideas that the audience is assumed to know, and basic propositions

that the audience is assumed to be satisfied with without further explanation. So the method of explanation of geometry is often applied outside of mathematics.

Closely related to the methods and ideas of explaining mathematical systems are the methods and ideas of symbolic logic. What is symbolic logic? Also called mathematical logic, this is what ordinary logic and reasoning developed into when the powerful symbolic methods of mathematics were applied during the last 100 years to the subject matter of logic and reasoning. The dictionary definition of symbolic logic is:

> a science of developing and representing logical principles by means of symbols in order to provide exact rules of deduction based on primitives (undefined concepts), postulates, and exact rules for the formation and transformation of expressions

How is symbolic logic related to explanation? One way is that symbolic logic makes completely precise the meaning of many frameworks for statements. For example, all of the following frameworks say exactly the same thing, and have the same meaning from the point of view of symbolic logic:

1) All ....s are ----s

2) ....s are ----s

3) Every .... is a ----

4) Each .... is a ----

5) A .... is a ----

6) The ....s are contained in the ----s

7) The class of .... is in the class of ----

8) ....ness is a sure sign of ----ness, etc.

This single common meaning is that one class of things is included in another class. The relation of inclusion is studied in symbolic logic in the same way as the relation of addition is studied in mathematics.

Accordingly, in order to explain well and to recognize precisely what relation needs to be expressed, it is useful to know something of symbolic logic. Symbolic logic identifies and defines many of the basic properties and relations used in all scientific and logical thinking.

Accordingly, it would be desirable for a person who has much explanation to write to have some knowledge of the technique for presenting mathematical systems, and if possible some knowledge of symbolic logic.

6. The Distinguishing Characteristics of Explanation

We are now ready to try to make clear how explanation is to be distinguished from:

-- instruction

-- exposition

-- technical writing

-- the presentation of mathematical systems

In contrast to instruction, explanation has to be more interesting and apply in many more situations. It applies in a great many fields where the relation of teacher and student does not exist or is very much changed.

Exposition is writing to convey information. Explanation is writing that makes some previously obscure topic clear and plain and understood. Explanation is therefore simply a part of exposition. A book of essays, a newspaper report, a speech in Congress, a resolution at a meeting, etc., are all exposition. But only here and there in this exposition will deliberate explanation be necessary.

In contrast to technical writing and the presentation of mathematical systems, explanation is more general, it applies in many more situations, and it hardly ever has a captive audience which is compelled or strongly motivated to understand. If a stranger on the street asks me for directions to an address in my neighborhood, he is not forced to understand the explanation I offer him: if I confuse him, he simply asks someone else.

As a result of these differences, it is reasonable to conclude that explanation is significantly different from closely related subjects and disciplines. It is therefore reasonable to treat explanation as a subject or discipline in its own right, and to seek to improve explanations by considering the differences between poor explanations and good ones.

Chapter 4

Factors Affecting the Quality of an Explanation

Outline

A. Factors Related to Subject Matter

1. Context
2. Ideas Assumed Known
3. Common Properties and Relations
4. Theses
5. Repetition
6. Examples
7. Development
8. Relations of Ideas to People
9. Self-Protection

B. Factors Related to Words

10. Spelling
11. Pronunciation
12. Wrong Words
13. Ambiguous Words
14. Fuzzy Words
15. Familiar or Strange Words
16. Frequency of Words
17. Length of Words
18. Connotation of Words
19. Vividness of Words
20. Word or Phrase

C. Factors Related to Sentences

21. Punctuation
22. Grammar
23. References of Words
24. Modifiers
25. Length of Sentences
26. Variety of Sentence Structure

D. Factors Related to Paragraphs

27. Paragraph Length
28. Unity within Paragraphs
29. Coherence within Paragraphs
30. Emphasis within Paragraphs

E. Factors Related to the Whole Explanation

31. Unity
32. Coherence
33. Emphasis
34. Brevity
35. Tempo
36. Accuracy
37. Simplicity
38. Style

F. Factors Related to the Audience

39. The Nature of the Audience
40. Importance
41. Interest
42. Entertainment
43. Motivation
44. Feedback
45. Effectiveness

G. Factors Related to the Explainer

46. Knowledge of the Subject Matter
47. Emotional Attitude
48. Sincerity
49. Skill
50. Time Available
51. Clerical Assistance
52. Assistance by Computer

In this chapter we survey a large number of factors affecting the quality of an explanation. In fact, we recognize 52 factors; and certainly a resourceful investigator could notice even more.

Are these too many to consider? I think not. Every now and then one of these factors becomes crucial to the success of an explanation; and it is a help to the explainer if he has recognized and thought about that factor ahead of time. As we said before, the first step to controlling some factor is to become aware that it exists.

Are 52 factors too many to hope to control? I think not, because many good editors (see the remarks by Philip Swain in Appendix 3) have discovered and do control all these factors. Then what passes through their editing becomes good explanation.

A. Factors Related to Subject Matter

The first group of factors in regard to the quality of an explanation are those related to subject matter, the ideas and relations which the explanation deals with, irrespective of the words and sentences that may be used.

1. Context. One of the first things an explainee needs to know is the context, the branch of knowledge, the general subject, of the thing being explained; for example, the context may be space travel, or stamp collecting, or customs at Eastertime in Russia, or any one of thousands of different subjects. Knowing this, the explainee can decide whether he is willing to listen or is not interested.

Poor Example: Newspaper headline: "JANIRO OR FLOOD NEXT FOR GREEN"

Guess 1: Two men named Janiro and Flood are next in line for a golf contest -- "green" referring to golf.

Guess 2: Two men named Janiro and Flood have as their next preference for some position a man named Green.

Right Answer, found out by accident months later: Janiro, Flood, and Green are boxers, and the next match for Green was Janiro or Flood

Good Example: Headline of an advertisement in a scientific magazine: "Do you want to control brightness .... measure light?" The advertisement went on with a discussion of instruments for this purpose.

The factor is the degree to which the subject or context is quickly and unmistakably identified.

2. Ideas Assumed Known. The next thing an explainee needs to know is the group of ideas assumed to be known and familiar to him. If the explainee finds that he does know this group of ideas, then he can read ahead with some confidence that he will

understand the explanation. Whether or not explainees already know the group of ideas assumed known, it is always a help to give identifying information for ideas mentioned.

Poor Example:

> Boolean algebra of course was largely the creation of George Boole, and 'The Laws of Thought' is a landmark in the development of the algebra of logic.

This example immediately raises but does not answer five questions: What is Boolean algebra? Who is George Boole? What is the "The Laws of Thought"? What is the algebra of logic? Is Boolean algebra the same as the algebra of logic? The sentence will be clear only if the reader already knows a great deal.

Better:

> Boolean algebra is a symbolic technique for calculating with AND, OR, NOT, and conditions. It was largely the creation of George Boole, the great English mathematician who lived 1815-1864. His book 'The Laws of Thought' published 1854 was a landmark in the development of Boolean algebra.

This example of explanation answers: What is Boolean algebra? Who is George Boole? What is "The Laws of Thought"? and avoids the other two questions, because it leaves out the term "algebra of logic", which in fact sometimes has a different meaning from Boolean algebra. It repeats the term "Boolean algebra", on the always useful principle of using the same term for the same idea.

Some writers of textbooks on English have maintained that it is monotonous and uninteresting to refer to the same idea always by the same expression, and they advocate variation of expression. This thesis is partly right and partly wrong; where the writer refers to almost the same idea, a different expression is a help to the explainee in making the difference clear; where the writer refers to exactly the same idea, the same expression is an essential help to the explainee in making the sameness clear.

Good Example:

> At the same moment another pirate grasped Hunter's musket by the muzzle, wrenched it from his hands, plucked it through the loophole, and with one stunning blow laid the poor fellow senseless on the floor. (Robert Louis Stevenson, "Treasure Island")

Poor:

> "At the same moment another pirate took Hunter's gun by the muzzle, took it from his hands, pulled it through the loophole, and with a hard blow knocked him senseless."

"Grasped" and "wrenched" report differences that "took" does not report.

The factor is: the extent to which the group of ideas assumed known are carefully and consistently selected and identified.

3. Common Properties and Relations. Often we can explain a strange idea by stating its properties and its relations to familiar ideas. It turns out from a study of all the ideas referred to by the common words of English (and other languages) that there is rather a short list of common properties and relations. The list of common properties and relations can be expressed in many different ways, but one form of it is given in Table 1.

A more compact list, worked out at a much earlier stage of English than the stage we are in now, is expressed in the six question-asking words of English:

Who? What? Why? When? Where? How?

If we try to explain a strange idea, the worth of our explanation can be largely measured by the extent to which we report its common properties and relations. If we omit some important properties and relations, our explanation is poor.

Sometimes, all the explanation of a strange idea that is needed may be found by scanning the checklist of Table 1, and specifying the common properties and relations that appear important.

Poor Example:

> "Plutonium is a new heavy element."

Table 1

COMMON PROPERTIES AND RELATIONS OF IDEAS -- CHECKLIST

Name, identification
Other names, repetition in other words, equivalents
Examples, instances
Definition, meaning, significance
Essence, theme, nature

Kind, sort, genus, species, class
Properties, nature, habits
Similar things, related things, associated things
posites, contrasts
Distinguishing characteristics

Things included in it, parts
Things of which it is a part
Context, environment, situation, field
Composition, material, substance
Structure, organization, construction

Activity, behavior, verb
Agents, doers, subject of verb
Products, object of verb, recipients
Manner, ways, adverbs
Size, dimensions, measurements

Quantity, number
Variation, range, average, deviations
Shape, form; solid, liquid, or gas
Weight, density
Appearance, look, color, luster

Sound, smell, taste, feel
Place, location, position, extent, prevalence
Time, duration, age, persistency
History, origin, causes, development
Future, results, effects, predictions

Purpose, function, use, worth, value
Advantages, disadvantages
Owners, users
Importance, relation to human affairs

Only three properties of plutonium are reported, and the words "new" and "element" are rather ambiguous.

Better Example:

> Plutonium is a heavy chemical element, recently identified by scientists and produced by the fission process from uranium. It is radioactive, and can be atomic fuel like uranium and thorium. It is very poisonous to human beings.

In this explanation, seven properties of plutonium are reported. Also, "new" has been replaced by "recently identified by scientists" and "element" has been replaced by "chemical element", so that neither expression is ambiguous any longer.

The factor is: extent to which the common properties and relations of a strange idea are all reported.

4. Theses. Often an explanation centers on a thesis, message, or proposition — and there may be several instead of one. These are the gist or meat of the explanation; they are what the explanation is to make clear. A thesis can usually be expressed by any one of a large number of statements, some poor, some good. A good statement of a thesis makes perfectly clear what idea we are talking about, and what assertion about the idea we are making.

Good Example:

> My thesis is not that with closer study of China we shall find little to criticize, but rather that through greater insight we shall be able to analyze China's problems better, and thus create an atmosphere in which we in our country can contribute better towards helping China solve her problems.

The factor is: the extent to which the gist of the explanation is made unmistakably clear.

5. Repetition. Often in an explanation, after an idea has been stated, even when stated as clearly as seems possible, it is good to repeat it in other words, to paraphrase it. This has two advantages: the explainee is helped to form the idea

or meaning in his mind, to some extent separated from the particular words that clothed it the first time; and the explainee has additional time to allow the idea to sink into his mind, to crease and groove it.

Example:

"to repeat the idea in other words, to paraphrase it."

The factor is: extent of repetition of key idea: using other words.

6. Examples. Examples, illustrations, instances, are essential in explaining. Over and over again we say "for instance?" or "give me an example", so that we can gain a clearer idea of what is being explained. And giving an example of something X which is a Y, is extremely useful proof of the falsehood of a statement like "No X's are Y's". We all form ideas in our minds by going from examples to supposed general rules.

Example:

The logarithm of a number is an exponent which put upon a base gives the number. For example, if the exponent 3 is put on the base 2, the result is 2^3, 2 to the 3rd power, or 2 times 2 times 2, which is 8: $2^3 = 8$. And so, 3 is the logarithm of 8 to the base 2: $3 = \log_2 8$.

The factor is: extent to which an explanation uses examples.

7. Development. In the course of an explanation, we need to present a group of ideas one after another. The sequence, arrangement, unfolding of the ideas constitute the "development" of the explanation. The understanding of a group of ideas is helped by good development.

The development of an explanation needs to be easy to understand. This means in most cases that the order to be used is an order based on one or more of the common properties and relations shown in Table 1. Such an order is often called a "natural" order, since it is an order connected with the nature of the idea.

A natural order may be based on relations, from simple to complicated; on distance, from nearby to farther away; on time, from early to late; etc.

Example:

The range of size of physical objects in the universe is equivalent to a

scale from one to 1,000,.....,000 where a total of 39 zeros have been placed after the numeral 1. One of the smallest physical objects is a proton, elementary particle of positive electricity. The largest physical item is the distance to the farthest observed galaxy or spiral nebula. The galaxy we are in is the Milky Way; its size is represented on the scale by 1 followed by 34 zeros. On this scale, a man is represented by 1 followed by 15 zeros.

Here the development is based on distance, but the sequence is: smallest distance; largest distance; then the distance belonging to the commonest example of a special term which could not be avoided (galaxy); and finally the distance belonging to one of the commonest of all items (man).

The factor is: quality and degree of natural, reasonable development.

8. Relations of Ideas to People. People are more interested in themselves and in other people than in anything else. This is both natural and true.

This statement is recognized in measuring the readability of writing. A standard factor in measuring readability is the proportion of "personal" references. What is "personal"? The references counted as personal are: the names of particular persons, like Alexander Graham Bell; personal pronouns, like "his, her, I, you"; masculine words like "nephew"; feminine words like "wife"; the two words "people" and "folks". References to persons of indefinite sex, like "student", or who may or may not be flesh and blood, like "buyer", and pronouns of indefinite reference, like "they" in "they say", are not counted as personal.

But a deeper and more significant personalizing of ideas consists of making clear the importance of the idea, of telling how closely it affects human beings and particularly you, the reader, the explainee. Advantages, disadvantages, uses, possibilities, values, purposes, results, applications -- these are the reasons why you give some idea attention, why you listen to some explanation. To make the importance of an idea clear is part of teaching, the part called "motivation", giving a student a motive for learning something; and nearly all of his motives rest on the

way the idea will affect him and his future.

Two of the persons concerned in explanation, of course, are the writer or explainer and the reader or the audience. It is often useful and sometimes necessary when explaining to mention the writer and reader. The best way is to use "you, I, we", almost as if you and I were talking, though avoiding some colloquial expressions that grate on people's feelings when seen on paper. Most textbooks nowadays are written using "you, I, we" all the way through. Old-fashioned expressions, "the student, the reader, the author, the undersigned" and other similar formal expressions are becoming more outdated all the time.

A still more valuable way of showing the relations of ideas to people is by illustrating the idea with a vivid scene, actors, and dialogue. Even if the action and talk are only barely convincing, the idea is often well explained. Rudolf Flesch says in "The Art of Readable Writing" (Harper's, 1949): "There's nothing on earth that cannot be told through a hero -- or heroine -- who is trying to solve a problem in spite of a series of obstacles. It's the classic formula."

Poor Example:

It is easy for one to despise what he cannot get.

Good Example:

One hot summer's day, a fox was strolling through an orchard till he came to a bunch of grapes just ripening on a vine which had been trained over a lofty branch. Just the thing to quench my thirst, quoth he. Drawing back a few paces, he took a run and a jump, and just missed the bunch. Turning round again, with a One, Two, Three, he jumped up, with no greater success. Again and again he tried after the tempting morsel, but at last he had to give it up, and walked away with his nose in the air saying: I am sure they are sour. (Aesop)

The factor is: extent of relatedness to the interests of people, and actual references to people.

9. Self-Protection. In an explanation, the writer may be expressing ideas which he is uncertain of or which are unpopular. Under conditions like these, terms intended to protect the writer creep into the explanation: "informed circles report", "White House spokesman", "some people think", "apparently", "alleged", "it is rumored that". An excess of caution has spoiled many an explanation.

A general hedge at the start of an explanation ought to be enough: "The information expressed in this report has not yet been completely verified, and we shall be grateful to any reader for his comments or criticisms."

The factor is: extent to which phrases protecting the writer do not hamper the explanation.

B. Factors Related to Words

The second group of factors we are concerned with are those related to words, the most important of all tools used by human beings for explanations.

10. Spelling. Logically, an explanation should not be much affected by so minor a detail as the spelling of the words contained in it. But practically, misspellings distract attention from the explanation, and therefore need to be avoided.

The factor is: absence of misspellings.

11. Pronunciation. Our ears are sensitive to sound patterns. In poetry, we expect rhymes, alliteration, and other types of musical sound patterns, and enjoy them. But explanation is prose, and sound patterns, such as occur in "of differing difficulty", are distracting, and need to be avoided.

The factor is: absence of conspicuous (and distracting) sound patterns.

12 to 15. Meaning. Many books can be written, and have been written, on the meanings of words. It is an immense subject. For purposes of a guide to explaining, however, the subject is narrower. Our problem is to select and arrange words so that the words will have only one meaning, and the explainee is very likely to understand that meaning. We shall assume that we know the meaning we want to convey.

In regard to meaning, we encounter essentially four kinds of problems: wrong words; ambiguous words; fuzzy words; and strange words.

12. Wrong Words. When we know the meaning we want to convey, the first word we think of quite often makes us say to ourselves "No, that's not it; that word does not fit." So we rack our brains, look in a dictionary or thesaurus, consider related words, and finally choose a word or a variation of phrasing that seems to fit our meaning.

But it is not always true that there is one right word. Sometimes English contains no single suitable way to express an idea, even a definite, well-known, and common idea. Then, depending on the circumstances, we are compelled to use different expressions for the same idea.

Example. A definite common idea that needs a single expression and does not have it is "or equivalently" or "or in other words". I often want to say "A, or equivalently B" when I wish to indicate that A and B are two terms both meaning the same idea that I am thinking of. And I mention B as close to A as possible so that you the reader will be as likely as possible to select a meaning which is common to A and B. Now the trouble with both the phrases "or equivalently" and "or in other words" is that they are long, and therefore have a tendency to be confusing in a short sentence which uses them:

We know the meaning, or equivalently idea, that we want to convey.

To express this relation, therefore, sometimes I use apposition, the two ideas written right next to each other and separated by commas:

We know the meaning, idea, that we want to convey.

Or I may use just plain "or" by itself, hoping that the context will show that I intend here that the meaning of "or" be "or equivalently" and not "or alternatively":

We know the meaning or idea that we want to convey.

Or I may put the second term in parentheses or between dashes:

We know the meaning (idea) that we want to convey.

We know the meaning -- idea -- that we want to convey.

Or I may use that rather convenient Latin abbreviation "i.e." meaning "id est", "that is":

We know the meaning, i.e., idea, that we want to convey.

If English were mathematics or symbolic logic, I could define and adopt a short word or symbol that would have the precise meaning "or equivalently". But since English does not provide for the invention of new words in this particular category, all we can do is hope that as English evolves over the next few centuries, usage will crystallize out and distinguish "or equivalently" from "or alternatively" just as "off" was thus distinguished from "of".

The factor is: absence of wrong words.

13. Ambiguous Words. Second, after we have put down words in an explanation, and have read them over to check them, we often notice possible other meanings that we did not intend: we find that our words are ambiguous. In real life our listener says to us, "Oh, I thought you meant ...."

Ambiguity happens because a great many words have many meanings, and different people will think of different meanings. A common word with many meanings, for example, is "line". In a large dictionary, you will find listed more than fifty.

The great determiner of the meaning of a word when used is the context, the association, in which it is used. And as soon as we detect an ambiguity, we change either the word or some part of its context, and generally succeed in removing the ambiguity. In fact, it is not very difficult to correct ambiguity provided you succeed in observing it.

The factor is: Absence of ambiguous meanings.

14. Fuzzy Words. Third is the problem of fuzzy words, words which do not convey clear ideas, but large collections of ideas all somewhat related. Some of them are long words and phrases, fairly abstract ones, like "democracy, facility, proviso, engineering principles, major inflationary advance". Others are short words which have a great many meanings like "take, get, draw".

Long fuzzy words and phrases are a disease in many types of writing, especially in government regulations, legal arguments, business writing, engineering papers, etc. But the disease is coming under control as a result of such books as Flesch's "The Art of Plain Talk" in the United States, and Ernest Gowers' "Plain Words" in England.

Short fuzzy words become a problem in Basic English, the vocabulary of about 1000 English words proposed around 1928 by the English scholar C. K. Ogden as an international language and for helping foreigners learn English. For, Basic English makes one of its greatest simplifications by reducing the number of verbs to a total of 18, which together with adverbs like "out" and prepositions like "up" produce many other needed meanings of verbs: "put up, put out, get up, get out, take up, take out".

Example:

> "Another question you will put is this: Why are all the banks not to be open again at the same time? The answer is simple, and I am certain that it will be clear to you: Your government is not ready to let the events of the last years take place again. We have no desire for, and will not have, another outburst of bank smashes." (A statement expressed in Basic English, from C. K. Ogden, "The System of Basic English")

Better:

> Another question you will ask is this: Why are all the banks not to be reopened at the same time? The answer is simple, and I know you will understand it: Your government does not intend that the history of the past few years shall be repeated. We do not want and will not have another epidemic of bank failures." (Speech of President F. D. Roosevelt, March 12, 1935)

As we read and compare these two examples (Ogden's being a translation of Roosevelt's), most of us will prefer "repeat" to "take place again", "epidemic" to "outburst", and "bank failure" to "bank smash".

Then there is the class of fuzzy words which are altogether clear and acceptable in ordinary writing, but which just don't convey enough meaning when you or I try to

carry out operations in the real world:

Poor Example:

"Bake the biscuits in a moderately hot oven, until they are light brown." Says the poor bride, "What in the world is moderately hot? and I don't have a window in my oven, and I can't tell when they become light brown."

Better:

"Bake them in an oven at 350 degrees for eighteen minutes."

In the field of indoor gardening and potting house plants, I still have not found out when something or other is "rich garden loam" and when it isn't.

The factor is: Minimum of fuzzy words, and maximum of words with meanings that are clear, accurate, and able to be applied.

15. Familiar or Strange Words. Fourth is the problem of strange words, words that are unlikely to convey meaning to the reader.

It is important when writing explanation to have a clear and vivid idea of the typical reader of that explanation, and to use words which will be familiar to him and not strange.

A word can have just the right meaning and yet be so strange and unusual that it can hardly be used in explanation.

Example:

"She was wearing a wimple."

A wimple is a covering around a woman's head and neck used in the time of Chaucer, and still used in the dress of some nuns. But this sentence will almost never convey meaning to a reader.

Such words and phrases are sometimes called "COIK" since they are "Clear Only If Known".

By using certain kinds of English constructions, however, we can mention a strange word in such a way that a reader can guess just about what it must mean.

Example:

"In poetry, we expect rhymes, alliteration, and other types of musical

sound patterns."

This kind of construction makes clear (at the least) that alliteration is a musical sound pattern expected to occur in poetry.

The factor is: minimum of words that are strange to the audience and used in constructions where their meaning cannot be guessed.

16. Frequency of Words. A good guide to the familiarity of a word is its frequency of occurrence. Many extensive counts of the occurrence of words have now been made. Several are reported in "The Teacher's Word Book of 30,000 Words" by E. L. Thorndike and Irving Lorge (New York: Teachers College, 1944). This book reports counts totaling about 18 million word occurrences, and tells how frequently each word in the first 30,000 occurred.

But for most purposes of explanation we already know enough to distinguish common and uncommon words on sight, and we can refer to the Word Book from time to time as we would a dictionary.

The factor is: large proportion of common words and small proportion of uncommon words.

17. Length of Words. Another good guide to the familiarity of a word is its length. The usual convenient measurement of length is number of syllables. In general, short, one-syllable, words are the most familiar, long, many-syllable words are the strangest.

In explanation, we can try consistently to use the shortest word for each idea. For example, "necessary" is four syllables"; "required" is two syllables; while "need" is one syllable, and has the added advantage of requiring us to mention the somebody who needs the something.

The factor is: large proportion of short words.

18. Connotation of Words. We are all aware of the connotation of words. A man usually speaks of his work as "my job". His personnel director usually speaks of a "position available". The president of the company is likely to speak of "the employment which our company provides". The dictionary carefully labels the connotations of many words

as "slang, colloquial, dialect, poetical," etc. Some four-lettered Anglo-Saxon words which nearly everyone knows cannot be written in print except by novelists who have become great successes.

The connotation of words not only reflects social class and attitudes but also triteness or freshness of expression. How fast? As fast as lightning. How cool? Cool as a cucumber. How happy? Happy as a lark. We think the speaker lazy, and without imagination.

Twenty years ago writing was much more formal than it is today. Nowadays the spoken language is much more used in writing. Readability experts tell business men "Write as you talk"; they even say "Write like you talk".

The idioms of spoken language of course are useful because they help people to explain in the way they know best. But your success with this maxim depends on your audience, the group of readers you desire to read what you write. For a wide group of readers, including people who will be distracted if you use "like" as a conjunction, we might phrase the maxim "Write the way you would talk carefully".

The factor is: extent of appropriate connotations of words.

19. Vividness of Words. Another scale on which words can be measured is whether they are vivid, picturable, active, concrete, or on the other hand colorless, unpicturable, passive, abstract. We can often translate ideas into vivid, picturable words. The presence of "is a fact, has been made, has taken place, has occurred" and the like is nearly always a sign that some verb has turned into a noun, and shouldn't have. The presence of "-tion, -ment, -ance", and quite a few other suffixes and prefixes are often signs of luscious ideas that have been thrown away in verbiage.

Example:

My detestation of parsnips is a long-standing and undeniable fact.

Better:

I have always hated parsnips.

The factor is: proportion of vivid, active, concrete words.

20. Word or Phrase. In order to refer to an idea, we often have a choice between using a single word or using a phrase made up of several words:

"The Decline and Fall of the Roman Empire" or "The Decline of the Roman Empire"

Doesn't "decline" include "fall"?

"The origin and development of ..." or "the development of ...":

Doesn't a discussion of "development" include a discussion of origin?

"I am writing you in regard to your letter of March 3" or "I am writing about your letter of March 3".

Doesn't the one word "about" do the work of "in regard to"?

People invented the phrase in order to express ideas more exactly, or to give greater prominence to an idea that might otherwise not receive sufficient attention. But people often use a phrase when they don't have to, and when it slows the quick perception of the idea in the reader's mind. And in some explanations, long phrases are used instead of single words so often that readers are distracted and annoyed.

Example:

"People invented the phrase in order to express ideas more exactly."

Less Clear:

"People invented the phrase to express ideas more exactly."

Here the grammatical structure of "to express ideas" is no longer clear; in this sentence it can be adjectival, modifying "phrase"; or it can be adverbial modifying "invented". So the single word "to" instead of the phrase "in order to" is here no gain.

Better:

"To express ideas more exactly, people invented the phrase."

The factor is absence of phrases where single words would serve as well.

C. Factors Related to Sentences

21. Punctuation. Punctuation is rather completely governed by rules; and the rules

are given in style-books, dictionaries, and textbooks on English. Conformity with the rules is a help in understanding an explanation, and departure from them is a distraction to the reader, and sometimes hides meaning.

The factor is: absence of mistakes in punctuation.

22. Grammar. Essentially, grammar consists of the rules of language for turning collections of words into sentences, and enabling them to assert a statement. Here again, the rules are rather completely fixed, and conformity with the rules is a help in giving information and receiving it.

In some areas, however, people today are changing grammar. A famous example is the split infinitive.

Example:

"He fails to completely understand grammar."

The adverb "completely" splits the infinitive "to understand". In this particular sentence there is no other good position for the word "completely". "He fails completely to understand grammar" and "He fails to understand grammar completely" are ambiguous, and "He fails to understand completely grammar" is a barbarous order of English words.

Schools used to teach the rule that you were not allowed to put any word between "to" and the infinitive verb following it. This is a silly rule, and authorities have shown that the rule was never true of English, and that it got into old grammar books by a mistake made through ignorance, and has been copied ever since by people who should have known better. But many people still believe the rule without questioning it; and if you are writing for an audience which includes such people, it is wise to take their belief into account, and not distract from your message by splitting an infinitive.

The factor is: absence of mistakes in grammar.

23. References of Words. A property of sentences which has a big effect on the clarity of an explanation is the references of words. We are all of us of course familiar with

the need for making a pronoun such as "it" refer to the last-mentioned noun. But all too easily we can fall into the trap of writing a sentence like:

Example:

> "For each person, the associations that cling to a word are very different, because actually meaning isn't in the word -- it is in the head of the person who hears it or uses it."

In this example, the word "it" occurs three times. But the reference of "it" is not systematic: at the first use, "it" refers to "meaning"; at the second and third uses "it" refers to "word". This kind of confusion of reference requires a reader to stop and mentally unscramble the sentence.

This is really a much bigger problem than just the problem of reference of the pronouns "I, you, we, he, she, it, they". Other words for which references need to be clear are "which, this, yes, no, to, do, so", as in "I would like to", "I do", "I should say so".

Example:

> "This is really a much bigger problem than just the problem of ..."

What does "this" refer to? It apparently refers to "the problem of references of words". And it is not clear whether the "this" clearly accomplishes that reference; it would be better to write "the problem of references of words is really a much bigger problem than just the problem of ..."

In fact, any word or phrase used a second time in a discussion may carry reference to the ideas associated with its first use.

Example:

> "Our ship steamed slowly up the bay, and the pilot came aboard from his small ship. Then the ship proceeded towards the mouth of the river."

In the second sentence, what does "the ship" refer to? Probably "our ship" -- but it could refer to the pilot's ship, which of course must follow along so as to take the pilot off again.

The factor is: minimum of ambiguous references.

24. Modifiers. In sentences we find two kinds of modifiers: words, phrases, or clauses, that modify the meaning of other words or phrases. One kind of modifier is restrictive, narrowing the meaning of the idea that is modified and producing a new, more limited idea. The other kind of modifier is nonrestrictive; it adds an idea, and really makes an additional assertion. You can't drop a restrictive modifier out of a sentence without completely altering the meaning. You can always drop a nonrestrictive modifier out of a sentence, and often improve the sentence thus making it sharper and clearer.

Example: (from the announcement by a large university of a two-week course on computing machines)

> "Versatile digital computers can quickly and reliably solve the constantly increasing problems posed by the large amounts of numerical and alphabetical information arising in scientific and engineering calculations, in business accounting and logistics problems, and in military control and strategical work."

This sentence, as is easy to see, is clogged with modifiers: "Constantly increasing" is a non-restrictive modifier, a remark thrown in and not necessary. "Numerical and alphabetical" is a non-restrictive modifier, a remark thrown in and not necessary. "Quickly and reliably": the "quickly" may be necessary; the "reliably" may not be necessary, -- for is a problem solved if it is not reliably solved? "Digital" is a necessary restrictive modifier. But the word "versatile" should apparently be replaced by the word "automatic", because it is necessary to exclude desk calculating machines operated by a clerk and to refer to automatic computing machines that run by themselves for hours at a time.

Restrictive modifiers are part of the tools for defining what we are talking about. How often do we need to repeat them? A good procedure generally is to take a fairly common word, give it some special meaning so as to refer to the defined idea,

and then use it consistently to refer to that idea.

Example:

> In the study of language, one of the technical terms is "phoneme", meaning "a significant sound of the words of the language". English has 43 phonemes by one count, such as the sound of "ng" heard in "sing" and "hung", the sound of "t" expressed by the "ed" in "crisped", etc.

Instead of using "phoneme" throughout this explanation, it is sufficient to use the word "sound", as a shortening of "significant sound of English".

Non-restrictive modifiers are essentially the subject matter of additional sentences. For the sake of variety of sentence structure, we may make one sentence out of two or three nonrestrictive modifiers. But we are ordinarily quite free to combine them or separate them as we like best.

The factor is: extent of good use of restrictive and nonrestrictive modifiers.

25. Length of Sentences. Just as the length of words is measured in syllables, so the length of sentences is measured in words. Both the length of words and the length of sentences are important readability factors. According to the studies of Rudolf Flesch, about 8 words is the average length of sentences in comic strips, and about 29 words is the average length of sentences in difficult technical writing, while an average of 17 words is correlated with good average writing that is easy to read.

This is only one half the story, however. How do we make long sentences shorter, and how do we make short sentences longer?

A sentence can easily be broken into two or more sentences. The chief breaking points to look for are words like "and", "or, "which". Even a long sentence of the form "if A, and B, and C, then D and F" can be broken into: "Suppose A. Suppose B, and C. Then D. And also E." In general any part of a sentence that seems to be too long can be broken off.

Example: The sentence in the previous example can be broken into two as follows:

> "Versatile digital computers can quickly and reliably solve the con-

stantly increasing problems posed by the large amount of numerical and alphabetical information arising in many areas. These areas include scientific and engineering calculations, business accounting and logistics problems, and military control and strategical work.

Combining short sentences into longer ones is discussed as "subordination" in textbooks on English.

The factor is: average length of sentence, around 17 to 20 words, and variety of sentence lengths.

26. Variety of Sentence Structure. Textbooks in English distinguish between sentences which are simple (one subject and one predicate), compound (two simple sentences joined by a coordinating conjunction), and complex (any sentence containing a subordinate clause). This distinction is one of the elements of variety in sentence structure. But there are many more elements of variety: presence or absence of lists of ideas in a sentence; contrasts between ideas; the location of modifiers; the choice of relations between ideas to be expressed; etc.

Examples:

A photographer making an exposure in dim light opens the iris of his camera.

If a photographer desires to make an exposure in dim light, he will open the iris of his camera.

A photographer willopen the iris of his camera in order to make an exposure in dim light.

A photographer opens the iris of his camera and thus prepares to make an exposure in dim light.

Variety of sentence structure affects explanation, because a reader is distracted, even annoyed, by monotony.

The factor is: presence of variety of sentence structure.

D. Factors Related to Paragraphs

The next larger unit of explanation after sentences is paragraphs.

27. Paragraph Length. The length of a paragraph is usually measured in number of sentences. Studies of readability show that the length of paragraphs has an effect on explanation. The longer the paragraph, after four or five sentences, the harder it is to read. The shortest a paragraph can be is one sentence; and although we have all seen excellent examples of paragraphs consisting of only one sentence, it would be a waste of the usual mental pause between paragraphs to make them all one sentence long.

The factor is: variety of paragraph length, usually between two and six sentences.

28. Unity within Paragraphs. It is rather hard for a sentence not to have unity, because every sentence is closely linked with the basic pattern of assertion: a subject, about which we are going to assert something, and a predicate, that which we assert about it. But it is rather easy for a paragraph not to have unity, because ideas must change from sentence to sentence in the paragraph, and yet there must be a central idea around which all the sentences in the paragraph cluster. Hence in paragraphs we have to work to accomplish unity.

The chief expression of unity in a paragraph is usually the "topic sentence", the sentence which expresses the central idea or topic of the paragraph. Sometimes it is at the beginning, sometimes in the middle, sometimes at the end, and sometimes implied and not written. The other sentences in the paragraph develop the central idea.

The factor is: presence of unity in paragraphs.

29. Coherence of Sentences. In a paragraph, the relations of sentences one after another makes an explanation either easy or hard to follow. Some of the relations between sentences are expressed externally either by conjunctions "and, since" or by connectives "for example, however". But the most important relation between sentences is the internal linking of ideas between them, which needs to be reasonable and clear. The connection of the ideas, the internal linking, is basic; and many writers who succeed in showing the internal linking leave out the "for example" and the "however".

Example: In the previous paragraph, the internal and external connections between the five sentences are:

(1) the topic sentence, about relations of sentences;

(2) external expression by connectives is one kind of relation; but

(3) internal expression by linking of ideas is another kind of relation;

(4) the internal linking is basic (a comment); and

(5) often leads to the omission of external connectives (another comment).

The factor is: degree of linking of ideas from sentence to sentence within paragraphs.

30. Emphasis within Paragraphs. In a paragraph, the first and last sentences have the most important positions. It is natural to give these positions to the most important sentences.

The factor is: extent to which sentences are appropriately emphasized within paragraphs.

E. Factors Related to the Whole Explanation

The next larger unit of explanation after paragraphs is generally the whole explanation. Sometimes with books, there are in-between units -- sections, chapters, parts, outlines, etc.

The factors of unity, coherence, and emphasis apply over again to the larger units of explanation, and are important. But there is not much advantage to discussing them over again. We shall just list them:

Factor 31: Degree to which the whole explanation possesses unity.

Factor 32: Degree to which the whole explanation is coherent, well-organized.

Factor 33: Degree to which there is appropriate emphasis in the whole explanation.

34. Brevity. An explanation, unlike some other kinds of writing, should be short. It is worth remembering that it should be "like a woman's skirt, long enough to cover the essential points, and short enough to be interesting."

Ways of shortening explanations include: tables, lists, diagrams, charts, pictures, working models, movies, sample practice situations, appendices, etc. You or I may at times not be able to use some of these devices, because they cost too much or we can't get hold of them; but we should always consider them. And it is always a good idea to remove any large quantity of data from the writing that people are expected to read, and put the data in tables or lists where those interested can refer to it.

The factor is: degree to which the explanation is no longer than needed.

35. Tempo. By the tempo of an explanation, we mean the rate at which we present important ideas in the course of the explanation. We can recognize compact explanation, where the important ideas are packed closely together; and loose explanation, where the important ideas have a great many easy words and ideas between them, padding. From general experience, we know that people have to be given time to absorb important ideas, and quite often, we give them the needed time by putting many easy ideas in between the ideas to be absorbed.

Example:

> In the previous paragraph, there are three main ideas to be absorbed: tempo; close-packed explanation; loosely-packed explanation. And the rest of the words are padding to give readers more time to take in the three main ideas.

The factor is: degree to which the tempo of the explanation is well adjusted to the readers.

36. Accuracy. The accuracy of an explanation is extremely important. This means: telling the truth carefully; choosing words that say no more than the truth, avoiding exaggeration; selecting the appropriate truth to be told; and reporting the truth in an unbiased way. Upon this factor depends the extent to which people can rely on the explanation.

The factor is: extent to which the explanation is accurate.

37. Simplicity. The simplicity of an explanation is of course also important. For

example, an explanation is likely to be simple when you choose simple and familiar examples to make your general statements clear. If I want to explain the idea "integral" from calculus, and can use as a familiar example water flowing into a bathtub, my explanation has an increased chance of being simple.

Probably nearly every facet of an explanation is open to the effort to make it simple.

The factor is: extent to which the explanation is simple.

38. Style. The meaning of "style" which I intend here is stated in the Shorter Oxford English Dictionary:

> "The manner of expression characteristic of a particular writer, or of a literary group or period; a writer's mode of expression, considered in regard to clearness, effectiveness, beauty, and the like. -- Those features of literary composition which belong to form and expression rather than to the substance of the thought or matter expressed."

Clearly, we have already covered some of the characteristics of writing included in style. Many other characteristics, including for example, the precise use from time to time of unusual words, smoothness of flow of words, etc., we have not covered. So let us place all of these characteristics into a final group of factors, which we shall call a single factor "style". We cannot easily measure them but we can notice that they exist.

The factor is: Presence of a pleasing style.

F. Factors Related to the Audience

39. The Nature of the Audience. Probably the most important of all the factors affecting an explanation is the nature of the audience. In other words, who are the people you are explaining to? What are their characteristics? Exactly the same explanation can be almost perfect for one audience and almost useless for another.

The nature of the audience, the characteristics of the audience, are: what they know, what they think about, what they are interested in, what words they understand,

what words they don't understand, what statements they will accept, what statements they won't accept, etc. All these audience characteristics influence greatly the explanation to be composed.

As Philip Swain says in his "Giving Power to Words" (see Appendix 3 to this report):

> The first rule of good writing is to know your reader -- who he is, how he thinks; what words he uses, what words he understands; what interests him, what bores him; wherein he is smart, wherein stupid.

And later he says:

> Note the many languages within our language. The college freshman learns that "the moment of force about any specified axis is the product of the force and the perpendicular distance from the axis to the line of action of the force." Viewing the same physical principle, the engineer says: "To lift a heavy weight with a lever, a man should apply his strength to the end of a long lever arm and work the weight on a short lever arm." Out on the factory floor the foreman shouts "Shove that brick up snug under the crowbar and get a good purchase; the crate is heavy." The salesman says "Why let your men kill themselves heaving those boxes all day long? The job's easy with this new long-handled pinch bar. With today's high wages you'll save the cost the first afternoon."

Furthermore, the audience consists of readers who are looking at their wrist watches. They are people who don't want to stay any longer than they have to or are inclined to.

The factor is: degree to which the explanation is adapted to the nature of the audience.

40. Importance. If something is really important to an audience, they will give it much attention.

No words at all are needed for one of the most important explanations that any

audience ever receives: a loud noise, a honk, a bell, is enough on the right occasion to tell an audience "look out, danger, you may be run over."

Even on less urgent occasions, important subjects poorly explained will receive much attention.

The factor: importance of the subject matter to the audience.

41. Interest. If something is really interesting to an audience, whether or not it is important, the audience will give it much attention. In May 1927 a U.S. air mail pilot crossed the Atlantic alone in an airplane, flying 3600 miles in 30½ hours from New York to Paris. Everybody became excited. On his return he was welcomed by the President of the United States and was raised in rank from Captain to Colonel Charles A. Lindbergh.

When an overwhelmingly interesting event happens, newspapers fill their pages with dozens of stories about various aspects. The explanations may not be very good. The stories may be poorly written and not worth reading in calmer moments. Yet the daily audience reads and reads, and is hungry for still more information.

Interesting subjects poorly explained will receive much attention.

The factor is: interest of the subject matter to the audience.

42. Entertainment. People like to laugh and be entertained; and many kinds of explanation can be highlighted with jokes and stories. A story judiciously chosen can produce an amazing impact on the minds of readers. Someone came running to tell Winston Churchill that Lindbergh had flown the Atlantic alone in a plane, and he said, "Well, I would be more surprised if a committee had done it". This story has stuck in my memory for more than 30 years. But too good a story produces a risk: the entertainment may be distracting; you may even get some readers who will skip the serious parts of your explanation, in the same way as the minds of TV listeners become more and more accustomed to ignoring the commercials.

The factor is: presence of an appropriate amount of entertainment.

43. Motivation. An explanation needs to give readers a motive for reading. If you

can make a reader really want to read, hardly anything will drag him away. If you fail to motivate him, his attention will wander in less than a minute.

Basically, the motive we appeal to is always the same. It consists of evidence of the close relation of the ideas in the explanation, to ideas that the readers are much concerned with. Motivation consists of making clear at the start and continually the intimate relation of the ideas offered to what people want.

The factor is: presence of motivation.

44. Feedback. Explanation should be produced for an audience that can freely register approval and disapproval, who can provide feedback. This is not true for many business writers who write regulations or advertisements, and many scientists who write papers or give addresses in scientific meetings. They have a captive audience or an audience that cannot be observed; so they produce many kinds of atrocious explanations. So, if you have a choice, you should produce your explanation in such a way that many readers can freely say they like it or they don't like it, and can provide information to you so you can adapt your explanation to them.

The factor is: presence of feedback from the audience.

45. Effectiveness -- Does It Work? Finally, an explanation can be tested for effectiveness on the audience: does it work? If the readers can do something they could not do before, or answer questions that they could not answer before, then the explanation has worked.

The factor is: degree of effectiveness of the explanation.

G. Factors Related to the Explainer

46. Knowledge of the Subject Matter. Every explanation depends on the knowledge that the explainer has of the subject he is seeking to explain. This knowledge needs to be broad and deep, broader and deeper than just the territory of the topic being explained. As an explainer you must be able to draw from much knowledge, experience, and thought, so that among all the possibilities of developing or clarifying a topic, you can aptly choose those aspects which fit best to your particular task at the moment. The

audience needs to have the feeling about you, "That guy knows his stuff."

The factor is: Broad and deep knowledge of the subject matter.

47. Emotional Attitude. An explanation also depends on the emotional attitude of the explainer towards the human beings in his audience. If you like them, are interested in them, look on them as your friends, and are eager to share your ideas with them and make them clear, then the audience acquires a very comfortable and friendly feeling about you as an explainer. And human beings are very quick to detect the opposite -- in the turn of a phrase, in the choice of an expression; they notice almost immediately condescension, or snobbishness, or impatience, or annoyance over the bother of having to explain, etc.; and they have a tendency to depart quickly from such explainers.

The factor is: presence of a warm emotional attitude by the explainer towards his audience.

48. Sincerity. Another factor in explanation is sincerity. Do you sincerely believe what you are explaining? If you use high-sounding words or false-sounding claims, or if you suggest that you are saying things that you don't yourself believe, you are likely to turn readers away from your explanation.

The factor is: presence of sincerity in the explanation.

49. Skill. An explanation depends on the skill of the explainer. Some of this skill comes from an awareness of the factors on which the effectiveness of an explanation depends. But even more comes from such things as: a good knowledge of human behavior; practice in expression and in writing; and practice in explaining where there is feedback.

The factor is: presence of skill in the explainer.

50. Time Available. An explanation requires a sufficient amount of time: time for the explainer to make the ideas clear, time for the audience to absorb them, also if possible, time for interaction between the audience and the explainer. Some explanations are doomed to failure, because there is inadequate time for them.

The factor is: sufficient time to give and absorb the explanation.

51 and 52. Assistance -- by Clerk and by Computer. An explainer can benefit greatly from certain kinds of assistance by intelligent clerks and by computer programs. This assistance can enable him to: keep track of the words he is using; make sure that each word not known to the audience is explained properly; make sure that examples and definitions are given; etc.

The factor is: assistance in the clerical details of making good explanation.

Chapter 5

The Classes of Words for Explaining

Outline

1. Words as Instruments for Explaining
2. Well-Known Words
3. Short Words
4. C. K. Ogden's "Basic English"
5. A Vocabulary for Explaining -- Another Approach
6. The Words for Explaining: Classes 1, 2, 3
7. Words That Do Not Need to be Known: Class 4
8. Phrases
9. The Procedure for Controlling Words for Explaining
10. An Example

1. Words as Instruments for Explaining

When we are preparing an explanation for an audience, words are the chief instruments we use. Of course many other devices -- pictures, charts, models, etc. -- are also used; but words are the prime tools: we do most of our explaining with them.

It has been said that "one good picture is worth a thousand words", and this is sometimes true. We may even assert that a working model of something is sometimes worth ten thousand words. But nevertheless, some words have to be given along with the pictures and the models, so that the audience can be guided in how to understand them. Furthermore, only a small number of subjects out of the enormous number of

possible subjects can be pictured or modeled; the others require words for their explanation.

Although words are the most important tools for explaining, they are not very good instruments. Like a stone arrowhead a word is a clumsy weapon. The main problem is that words mean different things on different occasions. The word "line", for example, has more than fifty meanings listed in a big dictionary. Which one applies in a given sentence? How shall we handle the puzzle of many meanings?

Basically, we solve the puzzle of many meanings by our long experience with language and our capacity to reason; the more experience we have had with language, the easier it is for us to decide what a word and the sentence containing it may mean. We seem to use an unconscious reasoning process: we notice how words are used together in patterns and we conclude what they must mean. Sometimes we even notice in our minds the lag until the words fit together and the sentence takes on meaning.

Clearly then, the more experience we have had with a word, the more likely we are to be able to use it, work with it, and understand it. And so explanation should be based chiefly on words with which we have had the most experience.

2. Well-Known Words

The first test then to apply to any word to be used in an explanation is, "Will the audience know it?" If the answer is yes, then we can use it without hesitating. If the answer is no, then we must either say the same thing in a different way, or decide that the word is one really needed, and use it giving additional explanation of it also.

This test is usually easy to apply. Most of the time it is easy to decide which words are well-known and which are not. For example, "telephone" is well-known to any audience, and "teletype" is not; "mechanical" is well-known to any audience, and "solid-state" is not.

Usually, also, it is rather easy to say what we want to in well-known words.

For we imagine ourselves talking out loud to a friend, and we find we leave out a lot of obscure words. In fact there are many nontechnical words and phrases which nearly everyone knows rather well. These words are a good resource for explanation.

3. Short Words

The second requirement for words used in explanations is that they should be short. People have had much more experience with short words than with long ones. "To explain in words of one syllable" implies explaining that is easy to understand. English has a great number of good short words; they carry a punch.

Probably the quickest and best measure of the shortness of a word is the number of its syllables. One-syllable words cannot be made shorter; so we may use them freely, if they are well-known. Words of several syllables need watching, and more watching as the number of syllables increases.

Often, if we find a long word that will explain something well, we can think of a shorter word that will explain it almost as well or even better. In such case we often make a little change in a sentence so as to use the shorter word. Take for example the word "necessary" -- four syllables. It is a well-known word but long; what can we use instead? The word "required" has two syllables; the word "needed", two. Now suppose we turn the sentence "X is necessary" around and say instead "we need X", or "you need X"; then we get down to words of one syllable. Or take for example "satisfactorily" -- six syllables. In talking, we would probably say "pretty well", but we do not like to write such colloquial talk; we can use instead "rather well", or "quite well" or "well enough". When we strip away long words in an explanation, we often find we have stripped away trash. It is much harder to be obscure with short words.

Sometimes though there is a flavor or precision to the long word that is missing with any short word near to it in meaning. Often in fact, such a long word has become well-known because of this virtue. If so, we shall not replace the long

word. Such a word is "efficiency" -- four syllables. Even some words of six syllables are well-known and have no good shorter equivalent, for example, "responsibility."

We come thus to the question "What is the set of words needed to explain simply a technical subject and produce a report which will be plain and clear to the audience?"

4. C. K. Ogden's "Basic English"

Students of language have studied the question, "What minimum vocabulary is needed so that we can say almost anything that we want to say?"

One of the important answers is one proposed by the English scholar C. K. Ogden. It is called Basic English and it is advocated for simple explanation and for teaching foreigners English. Basic English is essentially a set of 850 English words. It contains 600 nouns, 150 adjectives, and 100 "operations". The 100 "operations" include 18 verbs, 36 prepositions and conjunctions, 31 adverbs, and 15 pronouns.

This set of words is supplemented in several ways. First, the following rules apply:

(1) plurals in "s";

(2) derivatives in "er", "ing", "ed" from 300 nouns;

(3) adverbs in "ly" from qualifiers;

(4) degree with "more" and "most";

(5) questions by inversion with "do";

(6) operators and pronouns conjugate in full.

Thus in Basic English we have for example, from the word "heat" in the original list, "heater", "heating", and "heated". This adds to Basic English probably 1500 or 2000 words. Second, "measurement, numeral, currency, calendar and international terms are included in their English form." This adds to Basic English about 60 international words including "telephone, geography", 19 calendar words including "Tuesday, August," and 31 names of numbers, including "million, quarter, twice" -- perhaps altogether 130 to 150 words. Finally, special lists of 50 or 100 words are included in Basic

English for discussion in special fields. For example, the word "supply" appears in both the list for business and the list for general science.

It is remarkable that almost everything can be said in Basic English. Also, an idea expressed in ordinary words can often be replaced by a simpler equivalent in Basic English. This step often makes the idea sharp and clear, because of the very clever choice of words for Basic English.

For the purposes of explanation however the set of words in Basic English is certainly too small. First, of the 100 operators, only 18 are verbs: "come, get, give, go, keep, let, make, put, seem, take, be, do, have, say, see, send, may, will." Yet several thousand verbs are in ordinary English, and many of them are short, familiar, colorful, and very clear. The same is true of many other English words that are not verbs. So, in Basic English we lose some of the force and vividness of ordinary English. As we read passages written in Basic English, we fairly often have the feeling of looking through darkened glasses: much of what we expect to tell apart we cannot.

5. A Vocabulary for Explaining: A Different Approach

In this report a different approach is taken towards establishing a good vocabulary for explaining. The question we ask is "What is the best set of words to be used for explaining simply a technical subject?" The answer we propose here is:

1. Use short words.
2. Use well-known words.
3. In addition, use other words if needed, keeping track of them, explaining them if necessary, and allowing only those words to remain that are really needed on careful consideration.

How is this procedure to be implemented?

6. The Words for Explaining

We implement this procedure by paying close attention to the actual nature and properties of the words used in an explanation. In an explanation there will be a set of words used for explaining. These words are words (1) that are supposed to be known already or to be learned while reading, and (2) that are used as building blocks in later explanation and definitions. Suppose that we call these words the words for explaining.

We can usefully classify these words into three groups:

Class 1: Words that are so familiar that every reader will know all of them; for example, "is", "like", "must".

Class 2: Words not specially defined that are familiar, but perhaps some reader may not know some of them; for example, "alternative", "expression", "computer".

Class 3: Words that are not familiar, that many readers are not expected to know, and that are specially defined and explained in the course of the explanation.

Many of the words in Class 3 are the special terms of the subject.

In writing a report it is normal practice for a writer to keep track of the words in Class 3. They are the defined terms, the special terms, of his subject. Regularly he will put these words in the index at the back of his report, and give for each such word the number of the page where it is defined or explained or occurs in an important way in discussion. Sometimes he will also include a glossary and give their definitions carefully in the glossary. But what division should be made between the other two classes?

A practical and easy way to separate most words between the first and second classes is on the basis of number of syllables. We put words of one syllable -- if not specially defined -- in Class 1. Also, if a word becomes two syllables by the addition of one of the endings "-s, -es, -d, -ed, -ing", we put it in Class 1 (unless

the meaning is changed as in "I am looking for my glasses"); for these endings do not make a word any harder to understand. In addition, we put into Class 1:

(1) numbers; for example, "186,000, 3/10".

(2) places; for example, "Philadelphia".

(3) nations, organizations, people, etc.: "Swedish; Bell".

(4) years and dates; for example, "June, 1967".

Of course, not all of these words will be familiar to every reader (for example, "quadrillion"), but in the way they occur, they are definite and clear. When reading a book, we usually are not puzzled by such words, for we can tell from the context just about what they must mean. These words are not inventoried; they are accepted as clear.

All remaining words for explaining -- chiefly, words of two or more syllables and not specially defined -- are put in Class 2, and are to be inventoried during the writing of the explanation or the report.

Many Class 2 words of course will be entirely familiar to every reader; but it is useful to play safe by listing them and considering them. In fact, the chief virtue of the listing is that no hard words enter the report without notice; we have a guarantee that no hard word will suddenly be sprung on the reader like a trap.

This method was used when the book "Giant Brains or Machines that Think" was written in 1947-49. Fewer than 1800 different words were in the Class 2 list at the end of writing all thirteen chapters of the book.

We shall be referring over and over again to these classes of words used in explanation; so it is useful to choose descriptive names also,even if they are only partially accurate:

Class 1 -- The One-Syllable Vocabulary

Class 2 -- The Key Vocabulary

Class 3 -- The Special Vocabulary

7. Words That Do Not Need to be Known

Now there are more words in an explanation than "words for explaining". And so we shall do well to recognize:

Class 4: Words that do not need to be known or learned, and which are not used in later explanation and definitions.

These occur in the report in such a way that understanding them, though helpful, is in no way essential. One subdivision of Class 4 includes the name that appears just once in a report, as a kind of side remark by the way; for example, "an application such as calculus". Even if we understand no calculus, this fact diminishes hardly at all our understanding of the explanation. (Such a word may also appear in the index but it is not a word for explaining.)

Among the subdivisions of Class 4 are:

(1) words occurring only in quotations, where only the general drift of the quotation is needed for understanding the report;

(2) words or other expressions occurring in illustrative examples, which are not used in later explanations in the report;

(3) words occurring in the titles of books, articles, papers, etc., telling an interested reader where he can find more information; etc.

The descriptive name for Class 4 that we shall choose is "The Allusion Vocabulary".

8. Phrases as Instruments for Explaining

In addition to words, we also need phrases as instruments for explaining.

If the meaning of a phrase can be guessed from the meaning of its individual words, then we can treat it simply as a collection of separate words. This is true for phrases such as "with regard to", "traffic control system", etc.

If the meaning of a phrase cannot be guessed from the meaning of its individual words, then we have to treat the whole phrase as a single word, an integral unit. This is true for "put up with" or "differential analyzer" or "source language".

Let's name these "integral phrases". In scientific work, a very common way of specifying a new idea is to choose some descriptive phrase, define it specially, and use that integral phrase as the name of the new idea; for example, "black box", which is often neither a box nor black. Such descriptive phrases are equivalent to many-syllable words. Such phrases will go as a whole into Class 2 if they are sufficiently familiar, and will go as a whole into Class 3 if not familiar.

Such a phrase is felt by a speaker to be a single word. Words like "sometimes", "screwdriver", "maybe", show their origin to be a phrase felt as a single word. In some other languages than English, there is a great tendency to signal this kind of phrase by writing the constituent words together as a single word. In English there is no regular signal to show that a phrase is being used as a single unit, and is felt as a single word. This is a drawback. Usually we must simply burden our memory with the fact that the phrase as a whole is being used as a single word: for example, "differential analyzer". Usually we cannot replace the phrase by something simpler, since the phrase is a specific name. We may be stuck with a jawbreaker, and be driven to using an abbreviation like "EDP" for "electronic data processing".

In a report, each phrase with a special meaning, when it first appears, should be marked with quotation marks or underlined, and should be explained there or nearby. In addition, in the index all the phrases having special meaning should appear again, and next to each the page number where its explanation may be found.

Of course, there will also be phrases that do not need to be understood, in order for the explanation to be clear. These phrases will be classified in Class 4.

9. Procedure for Controlling Words for Explaining

Suppose we have an explanation to be improved. How shall we perform the procedure outlined above for controlling the vocabulary of words for explaining?

Here is the procedure:

(1) Make sure that the explanation is typed with double or triple line spacing.

(2) For each word (or integral phrase) found in the explanation, write underneath it:

- 3, if the word (or phrase) is a special term that needs to be explained to the assumed audience;
- 4, if the word (or phrase) does not actually have to be known in order to understand the explanation;
- 1, if the word has not previously been marked and has one syllable (after dropping any "-s, -es, -d, -ed, -ing");
- 2, if the word (or phrase) has not been previously marked and has two or more syllables.

(3) Collect in a list all the terms marked 3, and consider how the vocabulary can be clarified by:

- leaving out any term not needed at this stage;
- calling the same idea by the same name each time it is mentioned;
- defining clearly any term kept in the list; etc.

(4) Collect in a list all the words (and phrases) marked 2, and consider how this collection can be clarified by:

- replacing long words by short ones that say the same thing;
- replacing complicated words or phrases by simpler ones that are just as precise.

(5) Collect in a list all the words and phrases marked 4, and consider how this collection can be shortened and simplified.

(6) Apply these considerations to making a new draft of the explanation.

10. An Example

How does this procedure actually work in an explanation? Here is an example. Following are the first two sentences of a report called "Principles of Sampling" (see the full text in Appendix 1). Under each word in these sentences the appro-

priate number of its explanation class has been written.

Whether by biologists, sociologists, engineers, or chemists,
2 1 2 2 2 2

sampling is all too often taken far too lightly. In the early years
3 1 1 1 2 2 1 1 2 1 1 2 1

of the present century, it was not uncommon to measure the claws
1 1 2 2 1 1 1 2 1 2 1 1

and carapaces of 1000 crabs, or to count the number of veins in
1 4 1 1 1 1 1 1 1 2 1 1 1

each of 1000 leaves, and then to attach to the results the
1 1 1 1 1 1 1 2 1 1 2 1

"probable-error" which would have been appropriate had the 1000
3 1 1 1 1 2 1 1 1

crabs or the 1000 leaves been drawn-at-random from the population of
1 1 1 1 1 1 3 1 1 3 1

interest.
2

In this illustration of an explanation, there are four special terms (belonging in Class 3, the Special Vocabulary):

sampling

probable-error

drawn-at-random

population

There is one term that is strange and unusual, but does not have to be understood in this context (and as it belongs in Class 4, the Allusion Vocabulary):

carapaces

There are 18 words in the Key Vocabulary (Class 2):

appropriate	engineers	present
attach	interest	results
biologists	lightly	sociologists
century	measure	taken
chemists	number	uncommon
early	often	whether

None of these are unusual or difficult. All the rest of the words are in the One-Syllable Vocabulary (Class 1).

So most of the work needed to guarantee that the explanation is clear to the audience may be focused on the five words in Classes 3 and 4. For example, at the first occurrence "sampling", which is a special term, is treated as if everybody reading this passage already knew what it was; and this assumption is hardly justified for this audience, and some revision of wording should be considered. For a second example, "carapaces" could be quite easily changed to "shells"; to talk of the shells of crabs (even if they do not have true shells) is certainly more suited to this audience than to talk of their carapaces.

Chapter 6

Vocabulary Analysis With a Computer

In the preceding chapter we recognized four classes of words in an explanation:

Class 3 -- The Special Vocabulary

Class 4 -- The Allusion Vocabulary

Class 2 -- The Key Vocabulary

Class 1 -- The One-Syllable Vocabulary

In the sample explanation given, we typed the words with triple line spacing, and put under each word as a tag the number of its explanation class.

In order to analyze the explanation from the point of view of vocabulary, we next need to tabulate the words in each class 3, 4, and 2, putting them in alphabetical sequence, tally each occurrence of each word, total the tallies for each word, and total the entire number of words in each class and the whole explanation.

Ordinarily we do not need to analyze the words in class 1 individually, but can usually take them on faith. We can assume regularly that any difficulty in understanding an explanation will not come from Class 1 words. For Class 1, however, we would like to know the total number of different words and the total frequency; and in order to find out these figures, we might just as well tabulate the Class 1 words also.

An Illustration

As an illustration, the two sentences of sample explanation were:

> Whether by biologists, sociologists, engineers, or chemists, sampling is all too often taken far too lightly. In the early years of the present

century, it was not uncommon to measure the claws and carapaces of 1000 crabs, or to count the number of veins in each of 1000 leaves, and then to attach to the results the "probable error" which would have been appropriate had the 1000 crabs or the 1000 leaves been drawn at random from the population of interest.

The first step in clerical analysis is to put under each word the number of its vocabulary class; the result is shown at the end of the last chapter. The next step is to summarize it, and the result of this step is shown in Table 1.

The Amount of Work

The clerical work to do all this analysis on an explanation of full length is of course enormous, and largely impractical, even with shortcuts. Furthermore, writing and revising an explanation involves several drafts. Each draft will change the occurrence and the frequency of words in the vocabulary. Even for a two-page explanation (of 700 words say), revised three times, we would have four drafts, and we would have to tag, classify, and inventory 2800 words.

The Computer Program

But with a programmed computer the work becomes simple and practical. A computer program, here called the "Vocabulary Analysis Program," has been completed and will do this work. The program has been produced for the PDP-7 Computer made by Digital Equipment Corporation, and contains about 800 instructions in symbolic machine language. The program is stated (NOT explained) in Appendix 1 to this report.

The text of the explanation is typed on-line into the computer in the following form:

> Whether by biologists, sociologists, engineers or scientists, sampling 3 is all too often taken far too lightly. In the early years of the present century, it was not uncommon to measure the claws and carapaces 4 of 1000 crabs, or to count the number of veins in each of 1000 leaves, and then

Table 1

ANALYSIS OF VOCABULARY BY CLERK

Class 3 -- The Special Vocabulary

drawn at random	1
probable error	1
sampling	1

Number of Different Terms, 3

Total Frequency, 3

Class 4 -- The Allusion Vocabulary

carapaces	1

Number of Different Terms, 1

Total Frequency, 1

Class 2 -- The Key Vocabulary

appropriate	1	measure	1
attach	1	number	1
biologists	1	often	1
century	1	present	1
chemists	1	results	1
early	1	sociologists	1
engineers	1	taken	1
interest	1	uncommon	1
lightly	1	whether	1

Number of Different Terms, 18

Total Frequency, 18

Class 1 -- The One-Syllable Vocabulary

all	1	leaves	2
and	2	not	1
been	2	of	5
by	1	or	3
claws	1	the	9
count	1	then	1
crabs	2	to	4
each	1	too	2
far	1	veins	1
from	1	was	1
had	1	which	1
have	1	would	1
in	2	years	1
is	1	1000	4
it	1		

Number of Different Terms, 29

Total Frequency, 55

to attach to the results the "probable-error" 3, which would have been appropriate had, the 1000 crabs or the 1000 leaves been drawn-at-random 3 from the population 3 of interest.

Errors in typing can be easily corrected. Revisions can be easily made. Then the command to the computer:

F$$

produces the entire frequency distribution in the form shown in Table 2.

Almost all the time needed to produce the typed frequency distribution shown in Table 2 is used for operating the keys of the console typewriter. The program, for explanations up to 700 words in length, is almost entirely limited by the output typewriter speed, 17 characters typed a second.

As shown in the above example:

-- words to be treated as indivisible phrases are so marked in input by putting hyphens between the words;

-- only a few tags have to be specified (the three 3's and the one 4 in this case), since the computer program can assign almost all the other tags.

With efficiency like this, vocabulary analysis of explanations is changed from impractical to practical.

Tagging Words of Class 2 and Class 1

The computer program automatically tags correctly large numbers of words in Class 2 and Class 1. The program distinguishes between words of one syllable (after dropping "-s, -es, -d, -ed, -ing") and words of two or more syllables. The program works correctly in far more than 99% of all words occurring. The program uses a dozen or so rules based on spelling; and these rules distinguish between words in Class 1 and words in Class 2 in almost all cases.

If, when you are using the program, you notice a word which belongs in Class 2 although the program using its rules has put it in class 1 (like "nothing" or "during")

Table 2

ANALYSIS OF VOCABULARY BY COMPUTER

WHETHER BY BIOLOGISTS, SOCIOLOGISTS, ENGINEERS, OR CHEMISTS, SAMPLING 3 IS ALL TOO OFTEN TAKEN FAR TOO LIGHTLY. IN THE EARLY YEARS OF THE PRESENT CENTURY, IT WAS NOT UNCOMMON TO MEASURE THE CLAWS AND CARAPACES 4 OF 1000 CRABS, OR TO COUNT THE NUMBER OF VEINS IN EACH OF 1000 LEAVES AND THEN TO ATTACH TO THE RESULTS THE "PROBABLE-ERROR" 3 WHICH WOULD HAVE BEEN APPROPRIATE HAD THE 1000 CRABS OR THE 1000 LEAVES BEEN DRAWN-AT-RANDOM 3 FROM THE POPULATION 3 OF INTEREST.

CLASS 3
DRAWN-AT-RANDOM 1
POPULATION 1
PROBABLE-ERROR 1
SAMPLING 1
NU 4 FU 4

CLASS 4
CARAPACES 1
NU 1 FU 1

CLASS 2
APPROPRIATE 1
ATTACH 1
BIOLOGISTS 1
CENTURY 1
CHEMISTS 1
EARLY 1
ENGINEERS 1
INTEREST 1
LIGHTLY 1
MEASURE 1
NUMBER 1
OFTEN 1
PRESENT 1
RESULTS 1
SOCIOLOGISTS 1
TAKEN 1
UNCOMMON 1
WHETHER 1
NU 18 FU 18

CLASS 1
1000 4
ALL 1
AND 2
BEEN 2
BY 1
CLAWS 1
COUNT 1
CRABS 2
EACH 1
FAR 1
FROM 1
HAD 1
HAVE 1
IN 2
IS 1
IT 1
LEAVES 2
NOT 1
OF 5
OR 3
THE 9
THEN 1
TO 4
TOO 2
VEINS 1
WAS 1
WHICH 1
WOULD 1
YEARS 1
NU 29 FU 55

then you simply modify the stored text by putting the tag 2 next to any convenient occurrence of the word in the stored text:

nothing 2, during 2

If you notice a word which the computer classifies in Class 1 and which should instead go into Class 2 (like "beyond" or "preempt" -- "eyo" and "ee" are treated as single vowel combinations in the middle of a word), then you simply modify an occurrence of the word in stored text so that it reads:

beyond 2, preempt 2

Numbers like "1000" are automatically tagged 1. Names like "M. L. Adams" should be typed "M-L-Adams 1". The numbers "1, 2, 3, 4" in the text must be changed into "one, two, three, four"; otherwise the computer program will treat them as tags for the preceding words.

Tagging Class 3 and Class 4

There is no way at present of giving a computer program enough knowledge of a subject and enough knowledge of the English language, so that the computer can assign correctly the tags 3 and 4.

Instead, whenever the explainer tags a word in that way, the computer program accepts the tag. In addition, the computer program assigns that tag to every other occurrence of that word. So the explainer only needs to tag any such word once.

Other Tags

The tagging subroutines in the computer program also provide for 29 more optional tags, which need to be single letters B to Z or digits 5 to 9 and 0, <u>except</u> A, I, and O which are words. The optional tags do not interfere with the tags 1, 2, 3, 4. In any listing of the words in an explanation class with their frequency, these optional tags also appear. These tags could be used with such meanings as:

B Basic term, that the audience has to understand

E Terms that can be eliminated

P Probably understandable

U Undefined term

The computer program assumes that if any tag occurs in any occurrence of a word, it applies to every occurrence of that word.

Commands to the Computer Program

We shall now consider the commands to the computer program.

Insertion. When the computer program is loaded, and waiting for you to use it, the first thing you need to do is to give it some text to be analyzed. To do this:

1. Type the letter I (which is the command for INSERT).
2. Type the string of characters which you wish to insert, in the same way as if you were typing on an ordinary typewriter. What you type appears on the console paper, as well as going into the computer program.
3. Make sure that you do not type more than 1000 characters. At 60 characters to a line on the average, this would imply not more than 15 lines.
4. When you finish (usually it is good to finish with a CARRIAGE RETURN), type ALT MOD, ALT MOD. (This prints $$ at the end of your text.) Then the computer program gulps what you have typed and puts it in the text buffer; the program then issues a "carriage-return-line-feed" command to the typewriter; and then the program sits waiting once more for your next instruction.

Printing. You may wish to see what is the text buffer, what text is stored in the machine. If so, type

P, ALT MOD, ALT MOD.

Then the program prints out on the typewriter all that is contained in the text buffer.

Corrections. If as you are typing, you notice that some characters are wrong, type RUBOUT. This erases the last character typed. Typing RUBOUT once more erases the

preceding character, etc. As you type RUBOUT, the character prints again so that for example changing CORRECTOIN to CORRECTION will look like:

CORRECTOINNIOION

as you type in.

If you do not find an error until after all the typed-in text is in the text buffer, you can made use of the following procedure based on "search" and "pointer location".

For example, if you notice that CORRECTION has been misspelled CORRECTOIN, and you want to correct it, you give the command:

JSCORRECTOIN$T$$

This says (1) JUMP the pointer to the start of the text in the text buffer; (2) SEARCH for the string CORRECTOIN (starting with the character just after the S and ending with the character just before the ALT MOD ($ sign); (3) TYPE the line of text in which it occurs. The way it will appear is:

THIS IS HOW THE CORRECTOIN\ SHOULD BE MADE.

The back slash (\) shows where the pointer is.

Now we want to delete the last three characters, and insert three new characters. The command is:

-3DI ION$T$$

This says: (1) DELETE the three characters preceding the pointer; (2) INSERT the string ION (the next $ stands for ALT MOD, and has the meaning here "end of string"); (3) TYPE the line in which the pointer is. The result will be:

THIS IS HOW THE CORRECTION SHOULD BE MADE

Other commands for correcting text are also available.

Analysis. Once you are satisfied with the text in the buffer, and desire to obtain the FREQUENCY analysis, you type the command:

F$$

The analysis is then typed out in the sequence Class 3, Class 4, Class 2, and Class 1.

As you watch the words being typed, you may notice that something is wrong --

for example that a word in Class 2 really should be in Class 3. You can stop the typing by pressing the space bar; then you can correct the text in the text buffer by inserting "space, 3" after an occurrence of the word, which you can find with the SEARCH command. Then you can return to the analysis by using the command F$$.

If you wish to see only one of the classes, such as class 3, you can type:

3F$$

and only class 3 will be typed. (Wherever $ occurs in these pages, it stands for pressing the ALT MOD key. The dollar sign ($) is however an acceptable character in the text.)

A fuller explanation of this program, and a punched paper tape expressing the program, are available by writing to the author.

Chapter 7

Judging and Improving an Explanation

Outline

1. Purpose of this Chapter
2. Selection of a Sample Explanation for Judging and Improving
3. Judging and Improving is a Long Process
4. The Original Explanation
5. Vocabulary Analysis of the Original Explanation
6. The Regular Plan for Improvement of Any Explanation
7. Natural Questions
8. Answers to the Natural Questions
9. Producing the Revision
10. The Revised Explanation
11. Vocabulary Analysis of the Revised Explanation
12. Comments and Conclusions

1. Purpose of this Chapter

In the preceding chapters we have talked about:

-- the nature of explanation;

-- factors affecting the quality of an explanation;

-- the explanation classes for the vocabulary:

The Special Vocabulary, Class 3

The Allusion Vocabulary, Class 4

The Key Vocabulary, Class 2

The One Syllable Vocabulary, Class 1

-- a computer program for determining the frequency of each word in each vocabulary class in the explanation.

But so far we have not taken any particular sample explanation and applied the various ideas mentioned to that sample and shown just how those ideas can be used in a concrete case to improve that explanation to a considerable extent.

This is what we shall do now in this chapter.

2. Selection of a Sample Explanation for Judging and Improving

We shall select, as an example of explanation, the first page and a half of a certain technical book that has been published by MIT Press and that has sold more than 5000 copies called, "LISP 1.5 Programmers' Manual." The subject of this book is an interesting and important language for programming automatic computers which is called "LISP". The name comes from the first three letters of "LIST" and the first letter of "PROCESSING"; and the name refers to the fact that this programming language deals not only with numbers and letters, but also with lists of numbers and letters. These lists include statements, commands, instructions, theorems, equations, and in general any expression that can be made up out of symbols. Consequently, LISP is a very powerful language, for many kinds of analysis of ideas. LISP probably will become, along with FORTRAN, and COBOL, one of the important languages for programming computers.

This sample explanation is a reasonably good sample to select, for several reasons. First, as explained above, it deals with an important subject, instead of an unimportant subject. Second, the explanation we start with was judged to be publishable; in fact it was published, and has been widely read by persons who wanted to understand the subject. Third, the explanation shows plainly many common faults in ordinary technical explanation that is published or written today. And finally, the process of revising and improving this sample explanation will demonstrate how

the techniques we are discussing in this report can be employed to improve explanation.

3. Judging and Improving is a Long Process

We need to point out at the start one more fact: that describing in detail the actual analyzing and revising of an explanation is incredibly long-winded and tedious.

The process of revising a piece of writing usually takes place rapidly and invisibly in an author's mind, not slowly and visibly on a piece of paper. The author thinks over the way that some idea has been expressed on the piece of paper in front of him, and he decides to change the form of expression so as to make the idea clearer. So he inserts some words; he crosses out other words; he changes some more words; he marks a phrase to be moved from somewhere to somewhere else; he may even cut the piece of paper apart and glue the fragments together in another sequence; and so on, and so on.

A full description of the author's processing as he goes on and on, together with the statement of all the reasons why the author makes the changes that he does make, is very long and drawn-out. It is as tedious as would be the description of everything that goes on in the mind of a chessplayer as he plays a game of chess, evaluating each position and deciding what his next move will be, all through the game. When watching the playing of chess, we notice the player's pause and the wait from one move to the next, and then we see the resulting move. But we almost never watch the process of revising a piece of writing, and so we find it hard to imagine the pauses, the changes of thought, the process, when we see only the final finished product of writing.

So, because the process of revision of writing is very long-winded and full of decisions about details, a large number of pages are needed. Therefore we make extensive use of tables, lists, etc., all of which we call by one name "exhibits". The exhibits are given one after another at the end of this chapter.

4. The Original Explanation

The sample explanation chosen (see Exhibit 1) comes from pages 1 and 2 of the "LISP 1.5 Programmer's Manual" by John McCarthy and others, first published by The MIT Press, Cambridge, Mass., in 1962, and reprinted in 1966.

The explanation is stated in full in Exhibit 1; below is its beginning:

> The LISP language is designed primarily for symbolic data processing. It has been used for symbolic calculations in differential and integral calculus, electrical circuit theory, mathematical logic, and other fields of artificial intelligence. ....

The manual was written by persons thoroughly familiar with the version of LISP called "LISP 1.5" (presumably halfway between LISP 1 and LISP 2), and it is an essential reference for anyone using LISP.

5. Vocabulary Analysis of the Original Explanation

The first thing we want to do with the original explanation is analyze its vocabulary. For this purpose we give it to the computer program which was described in Chapter 6. The form in which it goes into the computer is given in Exhibit 2, and the output from the computer program is given in Exhibits 3A, 4A, 5A, and 6. For a number of reasons, the output is edited (see Exhibits 3B, 3C, 4B, and 5B). The main results of the vocabulary analysis are shown in Table 1. The classification of the words in the special vocabulary is shown in Table 2. The seven words in the Key Vocabulary which can be replaced by better words are shown in Table 3.

It is clear that the quality of this explanation is not very good.

- -- 34 special terms for 379 words of explanation is a very high proportion;
- -- 27 special terms occur only once, and they are logically candidates for removal;
- -- 7 special terms can be directly eliminated;
- -- 1 ambiguous term must be removed;

Table 1

VOCABULARY ANALYSIS OF THE ORIGINAL EXPLANATION --

SUMMARY

Explanation Class	Explanation Subclass	Description	Different Words		Occurrences	
			Number	Percent	Number	Percent
3		The special vocabulary	34	18.3	58	15.3
	P	Probably understandable	9			
	B	Basic terms	5			
	D	Defined terms	0			
	U	Undefined terms	12			
	E	Can be eliminated	7			
	M	Ambiguous terms	1			
4		The Allusion Vocabulary	20	10.8	20	5.3
2		The Key Vocabulary	62	33.3	74	19.5
	A	Acceptable words	55			
	E	Words that can be replaced by better ones	7			
	X	Unfamiliar, unacceptable words	0			
1		The One-Syllable Vocabulary	70	37.6	227	59.9
		Total	186	100.0	379	100.0

Table 2

THE ORIGINAL EXPLANATION --

THE SPECIAL VOCABULARY -- INVENTORY AND FREQUENCY

Subclass	Description	Special Term	Frequency
P	Probably understandable:		
		allocating storage	1
		available memory	1
		data	2
		execute (programs), execution (of a program)	2
		generate (programs)	1
		interpret (programs)	1
		machine language	1
		memory structure	1
		notation	1
B	Basic terms being explained:		
		atomic	1
		atomic symbol(s)	6
		LISP	4
		meta language	1
		S-expressions(s)	12
U	Undefined terms:		
		branching tree type of structure	1
		combine (two S-expressions)	1
		formal mathematical language	1
		form (S-expressions)	1
		isolated	1
		length (of an S-expression)	1
		list structures	1
		recursive (definition)	1
		recursive functions of S-expressions	2
		S-expression notation	1
		significant subexpressions	1
		source language	1
E	Terms that can be eliminated:		
		higher level languages	1
		LISP language	3
		LISP programming system	1
		split	1
		string	1
		symbolic calculations	1
		symbolic data processing	1
M	Ambiguous term:		
		symbolic expressions	1
		Total occurrences,	58

Table 3

THE ORIGINAL EXPLANATION --
THE KEY VOCABULARY -- WORDS THAT CAN BE REASONABLY REPLACED

Word	Frequency	Replacement
capable	1	able
concise	1	brief
distinct	1	different
elementary	1	simple
necessity	1	need
primarily	1	chiefly
readily	1	easily

-- 12 undefined terms should either be defined or eliminated;

-- 7 words in the key vocabulary could reasonably be replaced by better words.

6. The Regular Plan for Improvement of Any Explanation

In general the plan for improvement of any explanation is as follows:

Re the special vocabulary:

1. Basic terms: Clarify the explanation. Answer natural questions about them.

2. Terms probably understandable: Remove those not really needed. Make those that remain clearer and more understandable.

3. Terms that can be eliminated: Eliminate them by rephrasing.

4. Ambiguous terms: Remove the ambiguity.

5. Undefined, unguessable terms: Either explain them or remove them.

Re the allusion vocabulary:

6. Remove the allusions that are not really needed. Make those that remain clearer and more useful.

Re the key vocabulary:

7. Words that can be eliminated by commoner ones just as precise in meaning: Replace them.

There is nothing unusual or remarkable about this set of directions. What usually happens though is that an author is so busy thinking about his ideas and trying to express them that he forgets to apply these directions in his writing. Also, most authors are much more interested in their ideas than they are in conveying understanding to an audience; and so they do not feel any great urge to apply directions such as these.

No computer program to apply these directions is at present contemplated. Such a program, though not impossible, is still a long way in the future.

7. Natural Questions

In order to answer natural questions about the four basic terms or ideas being defined "LISP, S-expression, atomic symbol, atomic", let us consider what are "natural questions". A guide to these questions appears in the list of common properties and relations which was given in an earlier chapter, and is shown again as Exhibit 14. It will be reasonable to examine the items in this list one by one, and wherever a topic there suggested applies to the four basic ideas being defined, consider how to apply it.

In the following section the paragraphs are labeled with the topics that apply.

8. Answers to the Natural Questions

Name or Identification. The four basic terms have names which identify them.

Examples. How shall we give an example of the language LISP? It is hard to give an example unless we have access to a computer. Then we could show how LISP as a system on a computer actually works. However, we can give some examples of expressions written in LISP.

Examples of S-expressions and atomic symbols are given in Sample Explanation No. 1. Better and more interesting examples can be given, and should be given since an example is a fine opportunity to convey information.

Definition. It would take pages to define LISP as a computer programming language. We may as well not try to do that. Definitions of "S-expression" and "atomic symbol" were given.

It happens that the definition of atomic symbol in the original explanation is no longer true. Most LISP systems now allow atomic symbols to be made in more ways than with capital letters and digits, and also to contain an indefinite number of characters.

Species. LISP is a computer-programming language, i.e. a language for programming computers. It is problem-oriented, i.e., it is adapted to expressing problems to be solved. It is machine-independent, i.e., it depends hardly at all on the nature

of the computing machine; in other words, if you know how to use LISP on one computer, you can use it with just a few minutes' instruction on an entirely different computer.

An S-expression is a kind of symbolic expression, an expression composed of symbols which may or may not be digits. An atomic symbol is another kind of symbolic expression.

Properties. Several of the properties of LISP, of S-expression, and of atomic symbol were stated in the original explanation. It may be desirable to state some more of their properties.

Context or Field. The context or field of all these ideas is the field of special languages for programming computers to solve problems.

Composition and Structure. The original explanation does tell precisely the structure of atomic symbols and S-expressions. LISP is a structure for dealing with S-expressions on a computer, so as to solve problems.

Products. The product of LISP when used is answers to problems. So it would be good to show a problem, the use of LISP in it, and the answer. But there may not be room in the revised explanation to do this.

Quantity. The number of S-expressions and atomic symbols is indefinitely large. The number of "LISPs" is essentially one, although there are some variations in LISP on different computers. There is also an important difference between what is called Pure LISP and what is called Practical LISP. Pure LISP is very tedious but more elegant mathematically. Practical LISP contains a number of modifications which make LISP practical to use on a computer.

Origin. It would be helpful to state the origin and also the derivation of LISP, S-expression, and atomic symbol.

Importance and Use. It would also be helpful to say something about the importance of LISP. The sample explanation did mention some of the applications of LISP. There are now more applications.

Advantages and Disadvantages. It would be desirable to say something about both the advantages and disadvantages of LISP, even in an introduction, so as to satisfy

the needs and curiosity of the audience. Again, there may not be room to do this.

9. Producing the Revision

We shall now try to write a revised explanation. It should be an introduction to LISP for an audience of computer programmers who do not know anything about LISP and may have only a slight mathematical background. It should not make the errors in explanation pointed out earlier. It should try to cover the same basic points as covered in the sample explanation: giving definitions of LISP, S-expression, atomic symbol. It should answer at least some natural questions from the audience. After we have written the revised explanation, we should classify the vocabulary, and assess it thoroughly in the same way as the original explanation was assessed.

The details of producing the revised explanation, with comments here and there on the reasons for the particular changes made, are shown in Exhibit 7.

10. The Revised Explanation

The revised explanation is shown in Exhibit 8. The first few sentences are as follows:

> Among the new languages for programming computers is a remarkable one called LISP. The name comes from the first three letters of "LIST" and the first letter of "PROCESSING". The reason for the name is that LISP is particularly useful for working with lists, which may be lists of numbers, lists of names, lists of symbols, and other kinds of lists.
>
> ....

11. Vocabulary Analysis of the Revised Explanation

We must now judge the revised explanation. In order to analyze the vocabulary of the revised explanation we first put it into the computer program described in Chapter 6; the input is shown in Exhibit 9. The output from the computer program and the editing and summarizing of the output are shown in Exhibits 10 to 13.

The main results of the vocabulary analysis of the revised explanation are shown in Tables 4 and 5.

Table 4

VOCABULARY ANALYSIS OF THE REVISED EXPLANATION -- SUMMARY

Explanation Class	Explanation Subclass	Description	Different Words		Occurrences	
			Number	Percent	Number	Percent
3		The Special Vocabulary	17	6.7	94	12.1
	P	Probably understandable	7			
	B	Basic terms	3			
	D	Defined terms	7			
	U	Undefined terms	0			
	E	Can be eliminated	0			
	M	Ambiguous terms	0			
4		The Allusion Vocabulary	24	9.5	25	3.2
2		The Key Vocabulary	92	36.5	171	22.1
	A	Acceptable words	92			
	E	Words that can be replaced by better ones	0			
	X	Unfamiliar, unacceptable words	0			
1		The One-Syllable Vocabulary	119	47.3	485	62.6
		Total	252	100.0	775	100.0

Table 5

THE REVISED EXPLANATION -- THE SPECIAL VOCABULARY -- INVENTORY AND FREQUENCY

Subclass	Description	Special Term	Occurrences
P	Probably understandable		
		data	1
		execute (a program)	3
		generate (programs)	1
		implement(ed)	1
		implementation	1
		interpret (programs)	2
		machine language	2
B	Basic terms		
		atomic symbol(s)	13
		LISP	26
		S-expression(s)	21
D	Defined terms		
		dot	2
		dot notation	4
		list notation	3
		mathematical system	2
		NIL	2
		symbolic expression(s)	8
		( )	2
U	Undefined terms	NONE	0
E	Terms that can be eliminated	NONE	0
M	Ambiguous terms	NONE	0
		Total,	94

We can now judge the quality of the revised explanation:

-- 17 special terms for 701 words of explanation is rather a low proportion;

-- 4 special terms occur only once and are candidates for removal;

-- no special terms can be directly eliminated;

-- there are no ambiguous special terms;

-- there are no undefined or unguessable special terms;

-- there are no words in the key vocabulary that could reasonably be eliminated.

Furthermore, the three basic terms are each used many times, which is an advantage in learning them.

Therefore, on the basis of the classification of the words used, it is reasonable to consider the quality of this explanation good.

12. Comments and Conclusions

Clarity. The evidence given demonstrates that the revised explanation is clearer. All terms that occur in the revised explanation are either defined or probably understandable or definitely in the vocabulary of the assumed audience.

Length. Although twice as many words are used in the revision as in the original explanation, this is probably an advantage rather than a disadvantage.

Nevertheless, the revised explanation could be shortened by leaving out some explanation, such as the explanation of "list notation". Also the explanation could be shortened in other ways.

A main reason for the increase in length is putting in words needed to answer the natural questions that the audience would ask about the three important basic terms being defined.

Utility of the Computer Program. How useful is the computer program? Would it not be more practical to apply the principles of revising without using the computer program?

Of course, it is possible to do much the same revising and improving of an expla-

nation without the use of the computer program.

But the computer program has the following virtues:

-- It becomes very easy to notice and judge words occurring in the explanation;

-- Because of the search feature in the program, each occurrence of a word can be found;

-- The total number of times that a word occurs can be easily known, as a step in deciding whether to keep the word or exclude it;

-- It becomes very easy to identify and examine the special vocabulary;

-- As an explanation is revised, the computer program can keep step with the revision, and produce up-to-the-minute tabulations of words and their frequency.

Examining the lists of words produced by the computer program on preliminary runs quickly catches mistakes in classifying and tagging words. All in all, the computer program focusses a searchlight on the vocabulary being used.

Since the computer program tells how often words are used, the knowledge that a special term is used many times identifies it as a key idea. The knowledge that a special term is used only once is an invitation to eliminate it.

When the number of special terms in an explanation of 500 words length is cut from 30 to 15 there is a clear gain. Fewer special terms (3) used more often (60 times) lead to more understanding and more learning by the audience.

Amount of Work. Clearly, the amount of work indicated by the exhibits for this chapter seems enormous. But in actuality it is not enormous. It simply shows clearly the kind of work done by good editors of magazine articles, good writers of technical manuals, and other persons professionally concerned with presentation of information. The technique once learned takes much less time to carry out than it does to talk about it. It is something like the technique of a good tennis player -- much easier to do and do well, than it is to describe and teach it.

Effort. Is it worth making the effort to improve an explanation in this way?

Much of the effort can be delegated to a computer, which can inventory and count

the words in the explanation classes. This focuses the attention of the explainer on what he needs to do.

Also, much of the effort is inevitably made whenever an author or editor changes a first draft of a piece of writing into a second draft. The author or editor considers each term, phrase, and sentence, and tries to clarify and improve the expression of the thought.

Knowledge Necessary for Revising. What about the amount of knowledge needed to clarify and improve an explanation?

Only a person who has a thorough knowledge of the subject being explained is able to revise it adequately.

Isn't it rather difficult to revise and improve an explanation? Doesn't it require a good deal of expert ability?

Certainly it does require, besides a good deal of technical knowledge, knowledge of English, and some ability; and revising is often difficult. In fact, revising may be impossible if the person revising (the editor) does not have a good knowledge of the technical field.

In such a case however the editor can:

-- look up terms in a dictionary

-- read and study in the field;

-- consult persons who know a good deal about the field and find out from them what probably is meant by the text;

In some cases the editor can take the explanation back to the author and find out from him in many places just what he meant to say.

The Set of All the Key Vocabularies. The words Tagged 2 in the explanation, consist of:

> words of two or more syllables that are used for explaining, that most readers know or probably know, and that are not explained

These constitute the Key Vocabulary.

After dozens of explanations have been analyzed, all the words in all the Key Vocabularies can be assembled and studied. Many of these words can be revised and replaced with advantage; and possible replacements to be considered could be turned into another useful guide in making ideas clear and understood.

For example, most of the time (but not always) "often" is just as good as "frequently", "chiefly" is better than "primarily", "shorten" is better than "abbreviate", "just right" is better than "precisely correct", "place" is just as good as "situation", and so on. It is not desirable to distort or dilute any desired meaning. But it is silly to use a long word when a short word will do a better job. And a guide that translates from long words to short words would be helpful.

Once this happens, it should be possible to recognize and specify precisely the common relations that are being asserted about terms, and adopt preferred ways of expressing these relations. For example, "if .... , ----" is a preferred way for expressing conditions; this way is better and more modern than inversion as in "Had he told me, I would have done it."

In fact, many other important and interesting possibilities exist, as a result of fast, efficient classification of words used in vocabularies for explaining.

Exhibits

1. The Original Explanation

2. The Original Explanation Prepared for Input into the Computer Program

3A. The Special Vocabulary of the Original Explanation,
As Output by the Computer Program

3B. The Special Vocabulary of the Original Explanation Post-Edited, Summarized, and Classified

3C. Detailed Reasoning for Placing the Special Terms into the Subclasses

4A. The Allusion Vocabulary of the Original Explanation,
As Output by the Computer Program

4B. The Allusion Vocabulary Post-Edited and Summarized

5A. The Key Vocabulary as Output by the Computer Program

5B. The Key Vocabulary Post-Edited, Summarized, and Classified

6. The One-Syllable Vocabulary

7. Detailed Reasoning for Producing the Revised Explanation

8. The Revised Explanation

9. The Revised Explanation Prepared for Input into the Computer Program

10A. The Special Vocabulary of the Revised Explanation,
As Output by the Computer Program

10B. The Special Vocabulary, Post-Edited, Summarized and Classified

11A. The Allusion Vocabulary of the Revised Explanation,
As Output by the Computer Program

11B. The Allusion Vocabulary Post-Edited and Summarized

12A. The Key Vocabulary of the Revised Explanation,
As Output by the Computer Program

12B. The Key Vocabulary Post-Edited and Summarized

13. The One-Syllable Vocabulary of the Revised Explanation

14. Common Properties and Relations of Ideas -- Checklist

Exhibit I

THE ORIGINAL EXPLANATION

The LISP language is designed primarily for symbolic data processing. It has been used for symbolic calculations in differential and integral calculus, electrical circuit theory, mathematical logic, and other fields of artificial intelligence.

LISP is a formal mathematical language. It is therefore possible to give a concise yet complete description of it. Such is the purpose of this first section of the manual.

LISP differs from most programming languages in three important ways. The first way is in the nature of the data. In the LISP language, all data are in the form of symbolic expressions usually referred to as S-expressions. S-expressions are of indefinite length, and have a branching tree type of structure so that significant subexpressions can be readily isolated. In the LISP programming system, the bulk of available memory is used for sorting S-expressions in the form of list structures. This type of memory structure frees the programmer from the necessity of allocating storage for the different sections of his program.

The second important part of the LISP language is the source language itself which specifies in what way the S-expressions are to be processed. This consists of recursive functions of S-expressions. Since the notation for the writing of recursive functions of S-expressions is itself outside of the S-expression notation, it will be called the meta language.

Third, LISP can interpret and execute programs written in the form of S-expressions. Thus like machine language and unlike most other higher level languages it can be used to generate programs for further execution. The most elementary type of S-expression is the atomic symbol.

Definition: An atomic symbol is a string of 30 numerals and capital letters; the first character must be a letter.

Examples:

A

APPLE

PART2

A4B66XYZ2

These symbols are called atomic because they are taken as a whole and are not capable of being split within LISP into individual characters. Thus AB, CD, and ABCD have no relation to each other except that they are three distinct atomic symbols.

All S-expressions are built out of atomic symbols and the punctuation marks "(", ")", and ".". The basic operation of forming S-expressions is to combine two of them to produce a larger one. From the two atomic symbols A1 and A2 one can form the S-expression (A1 . A2).

Definition: An S-expression is either an atomic symbol or it is composed of these elements in the following order: a left parenthesis, an S-expression, a dot, an S-expression and a right parenthesis.

Note that this definition is recursive.

Examples: ATOM

(A . B)

(A . (B . C))

((A1 . A2) . B)

((U . V) . (X . Y))

((U . V) . (X . (Y . Z)))

Exhibit 2

THE ORIGINAL EXPLANATION PREPARED FOR INPUT

INTO THE COMPUTER PROGRAM

THE LISP-LANGUAGE 3 IS DESIGNED PRIMARILY FOR SYMBOLIC-DATA-PROCESSING 3 . IT HAS BEEN USED FOR SYMBOLIC-CALCULATIONS 3 IN DIFFERENTIAL-AND-INTEGRAL-CALCULUS 4 , ELECTRICAL-CIRCUIT-THEORY 4 , MATHEMATICAL-LOGIC 4 , AND OTHER FIELDS OF ARTIFICIAL-INTELLIGENCE 4.

LISP 3 IS A FORMAL-MATHEMATICAL-LANGUAGE 3 . IT IS THEREFORE POSSIBLE TO GIVE A CONCISE YET COMPLETE DESCRIPTION OF IT. SUCH IS THE PURPOSE OF THIS FIRST SECTION OF THE MANUAL.

LISP DIFFERS FROM MOST PROGRAMMING LANGUAGES IN THREE IMPORTANT WAYS. THE FIRST WAY IS IN THE NATURE OF THE DATA 3. IN THE LISP-LANGUAGE, ALL DATA ARE IN THE FORM OF SYMBOLIC-EXPRESSIONS 3 USUALLY REFERRED TO AS S-EXPRESSIONS 3. S-EXPRESSIONS ARE OF INDEFINITE LENGTH 3 AND HAVE A BRANCHING-TREE-TYPE-OF-STRUCTURE 3, SO THAT SIGNIFICANT-SUBEXPRESSIONS 3 CAN BE READILY ISOLATED 3 . IN THE LISP-PROGRAMMING-SYSTEM 3, THE BULK OF AVAILABLE-MEMORY 3 IS USED FOR STORING S-EXPRESSIONS IN THE FORM OF LIST-STRUCTURES 3. THIS TYPE OF MEMORY-STRUCTURE 3 FREES THE PROGRAMMER FROM THE NECESSITY OF ALLOCATING-STORAGE 3 FOR THE DIFFERENT SECTIONS OF HIS PROGRAM.

THE SECOND IMPORTANT PART OF THE LISP-LANGUAGE IS THE SOURCE-LANGUAGE 3 ITSELF WHICH SPECIFIES IN WHAT WAY THE S-EXPRESSIONS ARE TO BE PROCESSED. THIS CONSISTS OF RECURSIVE-FUNCTIONS-OF-S-EXPRESSIONS 3. SINCE THE NOTATION 3 FOR THE WRITING OF RECURSIVE-FUNCTIONS-OF-S-EXPRESSIONS IS ITSELF OUTSIDE OF THE S-EXPRESSION-NOTATION 3, IT WILL BE CALLED THE META-LANGUAGE 3.

THIRD, LISP CAN INTERPRET 3 AND EXECUTE 3 PROGRAMS WRITTEN IN THE FORM OF S-EXPRESSIONS. THUS LIKE MACHINE-LANGUAGE 3, AND UNLIKE MOST OTHER HIGHER-LEVEL-LANGUAGES 3, IT CAN BE USED TO GENERATE 3 PROGRAMS FOR FURTHER EXECUTION 3.

THE MOST ELEMENTARY TYPE OF S-EXPRESSION IS THE ATOMIC-SYMBOL 3. DEFINITION: AN ATOMIC-SYMBOL IS A STRING 3 OF NO MORE THAN 30 NUMERALS AND CAPITAL LETTERS; THE FIRST CHARACTER MUST BE A LETTER.

EXAMPLES: APPLE 4, PART2 4, A4866XYZ2 4 . THESE SYMBOLS ARE CALLED ATOMIC 3 BECAUSE THEY ARE TAKEN AS A WHOLE AND ARE NOT CAPABLE OF BEING SPLIT 3 WITHIN LISP INTO INDIVIDUAL CHARACTERS. THUS AB 4, CD 4, AND ABCD 4 HAVE NO RELATION TO EACH OTHER EXCEPT THAT THEY ARE THREE DISTINCT ATOMIC-SYMBOLS 3.

ALL S-EXPRESSIONS ARE BUILT OUT OF ATOMIC-SYMBOLS AND THE PUNCTUATION MARKS (,), AND . THE BASIC OPERATION FOR FORMING 3 S-EXPRESSIONS IS TO COMBINE 3 TWO OF THEM TO PRODUCE A LARGER ONE. FROM THE TWO ATOMIC-SYMBOLS A1 4 AND A2 4 ONE CAN FORM THE S-EXPRESSION 3 .... .

DEFINITION: AN S-EXPRESSION IS EITHER AN ATOMIC-SYMBOL 3 OR IT IS COMPOSED OF THESE ELEMENTS IN THE FOLLOWING ORDER: A LEFT PARENTHESIS, AN S-EXPRESSION, A DOT, AN S-EXPRESSION, AND A RIGHT PARENTHESIS.

NOTE THAT THIS DEFINITION IS RECURSIVE 3.

EXAMPLES: .... .

Exhibit 3 A

THE SPECIAL VOCABULARY OF THE ORIGINAL EXPLANATION,

AS OUTPUT BY THE COMPUTER PROGRAM

```
3F$$


CLASS 3
ALLOCATING-STORAGE 1
ATOMIC      1
ATOMIC-SYMBOL 3
ATOMIC-SYMBOLS 3
AVAILABLE-MEMORY 1
BRANCHING-TREE-TYPE-OF-STRUCTURE 1
COMBINE     1
DATA        2
EXECUTE     1
EXECUTION   1
FORMAL-MATHEMATICAL-LANGUAGE 1
FORMING     1
GENERATE    1
HIGHER-LEVEL-LANGUAGES 1
INTERPRET   1
ISOLATED    1
LENGTH      1
LISP        4
LISP-LANGUAGE 3
LISP-PROGRAMMING-SYSTEM 1
LIST-STRUCTURES 1
MACHINE-LANGUAGE 1
MEMORY-STRUCTURE 1
META-LANGUAGE 1
NOTATION    1
RECURSIVE   1
RECURSIVE-FUNCTIONS-OF-S-EXPRESSIONS 2
S-EXPRESSION 5
S-EXPRESSION-NOTATION 1
S-EXPRESSIONS 7
SIGNIFICANT-SUBEXPRESSIONS 1
SOURCE-LANGUAGE 1
SPLIT       1
STRING      1
SYMBOLIC-CALCULATIONS 1
SYMBOLIC-DATA-PROCESSING 1
SYMBOLIC-EXPRESSIONS 1
NU 37 FQ 58
```

Exhibit 3 B

THE SPECIAL VOCABULARY OF THE ORIGINAL EXPLANATION

POST-EDITED, SUMMARIZED, AND CLASSIFIED

Subclasses:

B	Basic term	U	Undefined	M	Ambiguous
P	Probably understandable	E	Can be eliminated		

Word	Freq.	Sub-class
allocating storage	1	P
available memory	1	P
atomic	1	B
atomic symbol(s)	6	B
branching tree type of structure	1	U
combine (two S-expressions)	1	U
data	2	P
execute (programs), execution (of a program)	2	P
formal mathematical language	1	U
form(ing) (S-expressions)	1	U
generate (programs)	1	P
higher level languages	1	E
interpret (programs)	1	P
isolated	1	U
length (of an S-expression)	1	U
LISP	4	B
LISP language	3	E
LISP programming system	1	E
list structures	1	U
machine language	1	P
memory structure	1	P
meta language	1	B
notation	1	P
recursive (definition)	1	U
recursive functions of S-expressions	2	U
S-expression(s)	12	B
S-expression notation	1	U
significant subexpressions	1	U
source language	1	U
split	1	E
string	1	E
symbolic calculations	1	E
symbolic data processing	1	E
symbolic expressions	1	M

Number, 34 Total Frequency, 58

Summary

B	Basic term	5
P	Probably understandable	9
U	Undefined	12
E	Can be eliminated	7
M	Ambiguous	1
		34

Exhibit 3 C

DETAILED REASONING FOR PLACING

THE SPECIAL TERMS INTO THE SUBCLASSES:

P	Probably understandable
B	Basic terms
U	Undefined
E	Can be eliminated
M	Ambiguous

As we look at the words in the special vocabulary and consider the assumed audience of programmers, we classify the following words as probably understandable to this audience:

allocating storage	interpret (programs)
available memory	machine language
data	memory structure
execute (a program)	notation
generate (a program)	

We mark these words P. For the other words, the process in detail of classifying them is as follows:

Paragraph 1:

(1) "LISP language". It seems clear that "LISP" and "LISP language" mean the same; the word "language" is here used as a kind of support to remind readers that LISP is a language. So we could write instead: "the language LISP." / In the table of the special vocabulary we can mark "LISP language" with the letter E, for "can be eliminated".

(2) "symbolic data processing". This might mean "processing of data consisting of symbols". But of course all data that a computer processes consists of symbols. So maybe this means: "data consisting of symbols as such rather than

numbers, that is, letters, words, and symbols as they occur in algebra, for example." So we could write instead:

"processing data consisting of symbols as in algebra". / E.

(This phrase still needs editing.)

(3) "symbolic calculations". All calculations use symbols, for numerals are symbols too; however, this must mean "calculations changing from one symbolic expression to another." / E.

Paragraph 2:

(1) "LISP" is a basic term being explained. The word "LISP" occurs 8 times. A person reading this explanation should gather at least 8 pieces of information from its 8 occurrences. / We can mark "LISP" B (for "basic").

(2) "formal mathematical language". Only a person who knows what a "formal mathematical language" is can benefit from this sentence. The assumed audience does not know. We will mark this term U for "undefined". If this idea is expressed in this explanation to this audience, then it should be defined. / U.

Paragraph 3:

(1) "symbolic expressions". This term occurs in the sentence:

In the LISP language all data occur in the form of symbolic expressions usually referred to as S-expressions.

This sentence is ambiguous. It might mean:

In LISP all data consist of symbolic expressions. The acceptable symbolic expressions in LISP are usually referred to as S-expressions.

(This by the way is true.)

Or it might mean:

In LISP all data consist of symbolic expressions. Symbolic expressions are usually referred to as S-expressions.

As a result, we are not able to decide on the meaning of "symbolic expressions," for it means either "expressions made up of symbols in general" or it means "expressions

that are made up of symbols and that are acceptable to LISP." We mark this term M for "ambiguous". / M.

(2) "S-expressions". This term occurs 11 times in the explanation. It is a basic term being explained. / We mark it B.

(3) "length (of an S-expression)". Up to this moment, the reader did not know that S-expressions had a length. / This term is marked U for "undefined".

(4) "branching tree type of structure". This term is clear enough so far as trees go. But how an expression made up of symbols can have this type of structure is not explained. (It is like the phrases "the feathers of the horses" or "square triangles".) / U.

(5) "significant subexpressions". How can a reader tell what is a "subexpression" and what is "significant"? No information given. Undefined. / U.

(6) "isolated". How do "subexpressions" get "isolated"? Undefined. / U.

(7) "LISP programming system". This is certainly the "programming system which puts LISP on a computer". This phrase can be eliminated. / E.

(8) "list structures". No information given. Not clear to an uninitiated programmer. / U.

Paragraph 4:

(1) "source language". No information given. No guesses fruitful. Undefined. / U.

(2) "in what way the S-expressions are to be processed". No explicit information has been given about the processing of S-expressions. Presumably, it means "changed from one form to another". / E.

(3) "recursive functions of S-expressions". "recursive" is a term somewhat familiar to mathematicians, but not to this audience. As for "functions of S-expressions", up to this point, nothing has been said about them. Undefined. / U.

(4) "S-expression notation". No information given about notation for S-expressions. Undefined. / U.

(5) "meta language". This phrase occurs in a sentence "since the notation ...., it will be called the meta language". It is not further used in this section of the explanation. It is apparently a basic term. / B.

Paragraph 5:

(1) "higher level languages". This term probably means "computer programming languages on a higher level than machine language," and so can be eliminated. / E.

Paragraph 6:

(1) "atomic symbol". The words "atomic" and "atomic symbol" occur a total of 7 times. It is a basic term being explained. / We mark it B.

(2) "string". With sufficiently astute guessing, this term can be understood in this context. It means a sequence of characters. At this occurrence, it could be eliminated by "a sequence of characters consisting of". / We mark it E.

(3) "split". This word occurs in the sentence:

> These symbols .... are not capable of being split within LISP into individual characters.

The interpretation of this might reasonably be:

> LISP is not able to determine that a given character is part of an atomic symbol using that character; for example, that the character P is a part of the atomic symbol APPLE.

(In general, a computer programming language can do anything of this nature that a person cares to program it to do. So the above interpretation is unreasonable. But it was true of LISP in 1962. It is no longer true of LISP because LISP has been extended with what are called "character-handling functions".) In any case, the word "split" in this unusual usage can be eliminated. / E.

Paragraph 7:

(1) "form" and "combine". These two words occur in the sentence:

> The basic operation for forming S-expressions is to combine two of them to produce a larger one.

What do "form" and "combine" mean? The next sentence shows an example, which seems simple to the point of being trivial. So how can this be the "basic operation"? Surely LISP exists for doing something useful with symbolic expressions, such as calculating one symbolic expression from another using certain rules. The basic operations of arithmetic are adding, subtracting, multiplying, and dividing. Surely the basic operations of LISP must be more than just "forming" and "combining" S-expressions. The meaning here is undefined. / U.

Paragraph 9:

(1) "recursive (definition)". Here again "recursive", though it is being illustrated, is not being defined. It is still undefined. / U.

This examination of the special vocabulary is summarized in Table 3B.

Exhibit 4 A

THE ALLUSION VOCABULARY AS OUTPUT BY THE COMPUTER PROGRAM

```
CLASS 4
A1              1
A2              1
A4B66XYZ2       1
AB              1
ABCD            1
APPLE           1
ARTIFICIAL-INTELLIGENCE 1
CD              1
DIFFERENTIAL-AND-INTEGRAL-CALCULUS 1
ELECTRICAL-CIRCUIT-THEORY 1
MATHEMATICAL-LOGIC 1
PART2           1
NU 12 FQ 12
```

Exhibit 4 B

THE ALLUSION VOCABULARY

AS POST-EDITED AND SUMMARIZED

differential and integral calculus	1
electrical circuit theory	1
mathematical logic	1
artificial intelligence	1
APPLE	1
PART2	1
A4B66XYZ2	1
A	1
ABCD	1
AB	1
CD	1
A1	1
A2	1
(A1 . A2)	1
ATOM	1
(A . B)	1
(A . (B . C))	1
((A1 . A2) . B)	1
((U . V) . (X . Y))	1
((U . V) . (X . (Y . Z)))	1
	20

Total Words, 20

Total Frequency, 20

Exhibit 5 A

THE KEY VOCABULARY

AS OUTPUT BY THE COMPUTER PROGRAM

CLASS	2
BASIC	1
BECAUSE	1
CAPABLE	1
CAPITAL	1
CHARACTER	1
CHARACTERS	1
COMPLETE	1
COMPOSED	1
CONCISE	1
CONSISTS	1
DEFINITION	3
DESCRIPTION	1
DESIGNED	1
DIFFERENT	1
DIFFERS	1
DISTINCT	1
EITHER	1
ELEMENTARY	1
ELEMENTS	1
EXAMPLES	2
EXCEPT	1
FOLLOWING	1
FURTHER	1
IMPORTANT	2
INDEFINITE	1
INDIVIDUAL	1
INTO	1
ITSELF	2
LANGUAGES	1
LARGER	1
LETTER	1
LETTERS	1
MANUAL	1
NATURE	1
NECESSITY	1
NUMERALS	1
OPERATION	1
ORDER	1
OTHER	3
OUTSIDE	1
PARENTHESIS	2
POSSIBLE	1
PRIMARILY	1
PROCESSED	1
PRODUCE	1
PROGRAM	1
PROGRAMMER	1
PROGRAMMING	1
PROGRAMS	2
PUNCTUATION	1
PURPOSE	1
READILY	1
REFERRED	1
RELATION	1
SECOND	1
SECTION	1
SECTIONS	1
SPECIFIES	1
SYMBOLS	1
TAKEN	1
THEREFORE	1
UNLIKE	1
USUALLY	1
WITHIN	1
WRITTEN	1

NU 65 FQ 74

Exhibit 5 B

THE KEY VOCABULARY

POSTEDITED, SUMMARIZED, AND CLASSIFIED

Subclass:

A Acceptable words

E Can be eliminated

X Unfamiliar, unacceptable words

For each word marked E, a suggested shorter or commoner synonym is shown in the Table.

Word	Frequency	Sub-Class	Alternative Word	Word	Frequency	Sub-Class	Alternative Word
basic	1	A		manual	1	A	
because	1	A		nature	1	A	
capable	1	E	(able)	necessity	1	E	(need)
capital	1	A		numerals	1	A	
character(s)	2	A		operation	1	A	
complete	1	A		order	1	A	
compose(d)	1	A		other	3	A	
concise	1	E	(brief)	outside	1	A	
consist(s)	1	A		parenthesis	2	A	
definition	3	A		possible	1	A	
description	1	A		primarily	1	E	(chiefly)
design(ed)	1	A		processed	1	A	
differ(s)	1	A		produce	1	A	
different	1	A		program(s)	3	A	
distinct	1	E	(different)	programmer	1	A	
either	1	A		programming	1	A	
element(s)	1	A		punctuation	1	A	
elementary	1	E	(simple)	purpose	1	A	
example(s)	1	A		readily	1	E	(easily)
except	1	A		referred	1	A	
following	1	A		relation	1	A	
further	1	A		second	1	A	
important	2	A		section(s)	2	A	
indefinite	1	A		specifies	1	A	
individual	1	A		symbol(s)	1	A	
into	1	A		taken	1	A	
itself	2	A		therefore	1	A	
language(s)	1	A		unlike	1	A	
larger	1	A		usually	1	A	
letter(s)	2	A		within	1	A	
				written	1	A	

Total Words, 61 Total Frequency, 73

Exhibit 6

THE ONE-SYLLABLE VOCABULARY

CLASS 1

30	1
A	10
ALL	2
AN	5
AND	11
ARE	6
AS	2
BE	5
BEEN	1
BEING	1
BUILT	1
BULK	1
CALLED	2
CAN	4
DOT	1
EACH	1
FIELDS	1
FIRST	3
FOR	7
FORM	4
FREES	1
FROM	3
GIVE	1
HAS	1
HAVE	2
HIS	1
IN	10
IS	14
IT	6
LEFT	1
LIKE	1
MARKS	1
MORE	1
MOST	3
MUST	1
NO	2
NOT	1
NOTE	1
OF	23
ONE	2
OR	1
OUT	1
PART	1
RIGHT	1
SINCE	1
SO	1
STORING	1
SUCH	1
THAN	1
THAT	3
THE	31
THEM	1
THESE	2
THEY	2
THIRD	1
THIS	4
THREE	2
THUS	2
TO	7
TWO	2
TYPE	2
USED	3
WAY	2
WAYS	1
WHAT	1
WHICH	1
WHOLE	1
WILL	1
WRITING	1
YET	1

NU 70 FO 227

Exhibit 7

DETAILED REASONING FOR PRODUCING THE REVISED EXPLANATION

Among the new languages for programming computers is a remarkable one called LISP. The name comes from the first three letters of "list" and the first letter of "processing". The reason for the name is that LISP is particularly useful for working with lists, which may be lists of numbers, lists of names, lists of symbols, and other kinds of lists. Lists classified in two or more ways are tables; an example is a train timetable, which is a list of the times that trains leave stations, classified in one way by train numbers and in a second way by stations on the railroad.

((This first paragraph mentions LISP, says what kind of thing it is, and explains where the name came from. Since lists are mentioned, the term "list" is clarified.))

The chief purpose of LISP is to process lists of symbols. The symbols may be numbers but more often are not numbers but symbols which express operations or names or words or other ideas. Symbols put together in a list or sequence having some meaning are called a symbolic expression.

((This is a revision of: "The LISP language is designed primarily for symbolic data processing."))

LISP was worked out in 1958-60 by a group of computer scientists at Mass. Inst. of Technology, Cambridge, Mass. The group included John McCarthy, Stephen B. Russell, Marvin L. Minsky, and others. They defined LISP in machine language for the IBM 7090 computer; that is, they "implemented" it for that computer, so that LISP would run and could be used.

((Origin of LISP, and introduction of the term "implemented".))

LISP has been used for calculations with symbolic expressions in differential and integral calculus, the theory of electrical circuits, mathematical logic, the playing of games of strategy, the design and testing of computer programs, and in

other fields where calculation with numbers is less important than calculation with symbols.

((Applications of LISP.))

Not only is LISP a language for programming computers but it is also a mathematical system, in the same way as the geometry of Euclid or the ordinary algebra of numbers is a mathematical system.

((Paraphrase of: "LISP is a formal mathematical language".))

((Omission of: "It is therefore possible to give a concise yet complete description of it. Such is the purpose of this first section of the manual." Reasons: (1) The manual did not actually give a concise yet complete description of LISP. (2) Even if an entity is not a "formal mathematical system", it would still be possible to give a concise yet complete description of the entity.))

LISP differs from most programming languages in several important ways. One difference is that all data that LISP deals with consists of symbolic expressions. The symbolic expressions that are acceptable to LISP are called "S-expressions."

A second difference is this: When programming in LISP, we do not have to allocate portions of the memory of the computer for storing symbolic expressions; this is done automatically by the implementation of LISP on the computer.

((Revision of Paragraph 3. The information omitted is not needed in the introduction.))

((Omission of paragraph 4. Reasons: It is not true that the "source language" (also called the "meta language") is necessary for the writing of "recursive functions of S-expressions". Also it is not true that all "source language" consists of "recursive functions of S-expressions."))

A third important difference is that LISP can interpret and execute programs

written in LISP. Thus, like machine language and unlike many other programming languages, LISP can be used to generate new programs. In addition, the way in which LISP interprets and executes LISP can be written in LISP.

((Revision of Paragraph 5.))

As we said above, symbolic expressions acceptable to LISP are called S-expressions. (The letter S in "S-expression" comes from the first letter of the word "symbolic".) What makes them acceptable?

The simplest kind of symbolic expression acceptable to LISP is one which regularly does not have any parts that LISP can recognize and that is treated by LISP as a unit. An acceptable symbolic expression of this kind is called an atomic symbol.

In almost all versions of LISP, capital letters written together without spaces (in the same way as letters are written together to make words in ordinary English) are treated as atomic symbols; for example, PRINT, PLUS, NULL. Also capital letters and digits written together in a group in which the first character is a capital letter is an atomic symbol; for example, PRIN1. In many versions of LISP, numbers are treated as atomic symbols of a special kind, since they have some properties that most atomic symbols do not have; for example, 22, 395. There is a unique atomic symbol represented as NIL or ( ), a pair of parentheses with nothing inside. It stands for an empty list, a list with no elements.

Examples of atomic symbols are:

PLUS	( )
22	NIL
A	D
M4	TRANSITIVERELATION

All S-expressions are either atomic symbols or they are built up out of atomic symbols in specified ways.

One way is this: Put together a left parenthesis, then an atomic symbol or S-expression, then a dot, then an atomic symbol or S-expression, then a right

parenthesis, and the result is an S-expression. This kind of S-expression uses what is called dot notation. The dot stands for an operator of LISP which is called CONS (pronounced "conss"). Examples of S-expressions in dot notation are:

(A . 22)

(B . (A . C))

(2 . (4 . (6 . ( ))))

((2 . 22) . ((3 . 33) . ((4 . 44))))

((A . B) . C)

(A . NIL)

(A . (B . NIL))

Since "S-expression" includes "atomic symbol", "an atomic symbol or S-expression" is the same as "an S-expression".

A second way is this: Put together a left parenthesis, an S-expression, a space, an S-expression, a space, an S-expression, ..., a space, an S-expression and a right parenthesis; and the result is an S-expression. This kind of S-expression uses what is called list notation. Any such expression is equivalent to an expression written in dot notation according to certain rules which will be explained later.

Examples of S-expressions in list notation are:

(A B C D E)

(PLUS 22 33 55)

(REVERSEOF (S U L P))

((W (K 50) (P 51)) (B (K 43) (P 54)))

((2 3 6) (2 4 8)) (3 5 15) (4 7 20))

All S-expressions can be written in dot notation. Not all S-expressions can be written in list notation.

((Revision of paragraphs 6, 7, and 8.))

Exhibit 8

THE REVISED EXPLANATION

Among the new languages for programming computers is a remarkable one called LISP. The name comes from the first three letters of "list" and the first letter of "processing". The reason for the name is that LISP is particularly useful for working with lists, which may be lists of numbers, lists of names, lists of symbols, and other kinds of lists. Lists classified in two or more ways are tables; an example is a train timetable, which is a list of the times that trains leave stations, classified in one way by train numbers and in a second way by stations on the railroad.

The chief purpose of LISP is to process lists of symbols. The symbols may be numbers but more often are not numbers but symbols which express operations or names or words or other ideas. Symbols put together in a list or sequence having some meaning are called symbolic expressions.

LISP was worked out in 1958-60 by a group of computer scientists at Mass. Inst. of Technology, Cambridge, Mass. The group included John McCarthy, Stephen B. Russell, Marvin L. Minsky, and others. They defined LISP in machine language for the IBM 7090 computer, that is, they "implemented" it for that computer, so that LISP would run and could be used.

LISP has been used for calculations with symbolic expressions in differential and integral calculus, the theory of electrical circuits, mathematical logic, the playing of games of strategy, the design and testing of computer programs, and in other fields where calculation with numbers is less important than calculation with symbols. Not only is LISP a language for programming computers but it is also a mathematical system, in the same way as the geometry of Euclid or the ordinary algebra of numbers is a mathematical system.

LISP differs from most programming languages in several important ways. One difference is that all data that LISP deals with consists of symbolic expressions.

The symbolic expressions that are acceptable to LISP are called "S-expressions".

A second difference is this: When programming in LISP, we do not have to allocate portions of the memory of the computer for storing symbolic expressions; this is done automatically by the implementation of LISP on the computer.

A third important difference is that LISP can interpret and execute programs written in LISP. Thus, like machine language and unlike many other programming languages, LISP can be used to generate and execute new programs. In addition, the way in which LISP interprets and executes LISP can be written in LISP.

As we said above, symbolic expressions acceptable to LISP are called S-expressions. (The letter S in "S-expression" comes from the first letter of the word "symbolic"). What makes them acceptable?

The simplest kind of symbolic expression acceptable to LISP is one which regularly does not have any parts recognizable by LISP and is treated by LISP as a unit. An acceptable symbolic expression of this kind is called an atomic symbol.

In almost all versions of LISP, capital letters written together without spaces (in the same way as letters are written together to make words in ordinary English) are treated as atomic symbols; for example, PRINT, PLUS, NULL. Also capital letters and digits written together in a group in which the first character is a capital letter is an atomic symbol; for example, PRIN1. In many versions of LISP, numbers are treated as atomic symbols of a special kind, since they have some properties that most atomic symbols do not have; for example, 22, 395. There is a unique atomic symbol represented as NIL or ( ), a pair of parentheses with nothing inside. It stands for an empty list, a list with no elements.

Examples of atomic symbols are:

PLUS	( )
22	NIL
A	D
M4	TRANSITIVE RELATION

All S-expressions are either atomic symbols or they are built up out of atomic symbols in specified ways.

One way is this: Put together a left parenthesis, then an atomic symbol or S-expression, then a dot, then an atomic symbol or S-expression, then a right parenthesis, and the result is an S-expression. This kind of S-expression uses what is called dot notation The dot stands for an operator of LISP which is called CONS (pronounced "conss"). Examples of S-expressions in dot notation are:

(A . 22)

(B . (A . C))

(2 . (4 . (6 . ( ))))

((2 . 22) . ((3 . 33) . ((4 . 44))))

((A . B) . C)

(A . NIL)

(A . (B . NIL))

Since "S-expression" includes "atomic symbol", "an atomic symbol or S-expression" is the same as "an S-expression".

A second way is this: Put together a left parenthesis, an S-expression, a space, an S-expression, a space, an S-expression, ... , a space, an S-expression and a right parenthesis, and the result is an S-expression. This kind of S-expression uses what is called list notation. Any such expression is equivalent to an expression written in dot notation according to certain rules which will be explained later.

Examples of S-expressions in list notation are:

(A B C D E)

(B)

(PLUS 22 33 55)

(REVERSEOF (S L U P))

((W (K 50) (P 51)) (B (K 43) (P 54)))

((2 3 6) (2 4 8) (3 5 15) (4 7 20))

All S-expressions can be written in dot notation. Not all S-expressions can be written in list notation.

Exhibit 9

THE REVISED EXPLANATION PREPARED FOR INPUT INTO THE COMPUTER PROGRAM

AMONG THE NEW LANGUAGES FOR PROGRAMMING COMPUTERS IS A REMARKABLE ONE CALLED LISP 3. THE NAME COMES FROM THE FIRST THREE LETTERS OF "LIST" AND THE FIRST LETTER OF "PROCESSING". THE REASON FOR THE NAME IS THAT LISP IS PARTICULARLY USEFUL FOR WORKING WITH LISTS, WHICH MAY BE LISTS OF NUMBERS, LISTS OF NAMES, LISTS OF SYMBOLS, AND OTHER KINDS OF LISTS. LISTS CLASSIFIED IN TWO OR MORE WAYS ARE TABLES; AN EXAMPLE IS A TRAIN TIMETABLE, WHICH IS A LIST OF THE TIMES THAT TRAINS LEAVE STATIONS, CLASSIFIED IN ONE WAY BY TRAIN NUMBERS AND IN A SECOND WAY BY STATIONS ON THE RAILROAD.

THE CHIEF PURPOSE OF LISP IS TO PROCESS LISTS OF SYMBOLS. THE SYMBOLS MAY BE NUMBERS BUT MORE OFTEN ARE NOT NUMBERS BUT SYMBOLS WHICH EXPRESS OPERATIONS OR NAMES OR WORDS OR OTHER IDEAS. SYMBOLS PUT TOGETHER IN A LIST OR SEQUENCE HAVING SOME MEANING ARE CALLED SYMBOLIC-EXPRESSIONS 3.

LISP WAS WORKED OUT IN 1958-60 BY A GROUP OF COMPUTER SCIENTISTS AT MIT-CAMBRIDGE-MASS 1. THE GROUP INCLUDED JOHN-MCCARTHY 1, S-B-RUSSELL 1, M-L-MINSKY 1, AND OTHERS. THEY DEFINED LISP IN MACHINE-LANGUAGE 3 FOR THE IBM-7090 4 COMPUTER, THAT IS, THEY "IMPLEMENTED" 3 IT FOR THAT COMPUTER, SO THAT LISP WOULD RUN AND COULD BE USED.

LISP HAS BEEN USED FOR CALCULATIONS WITH SYMBOLIC-EXPRESSIONS IN DIFFERENTIAL-AND-INTEGRAL-CALCULUS 4, THE THEORY-OF-ELECTRICAL-CIRCUITS 4, MATHEMATICAL-LOGIC 4, THE PLAYING OF GAMES OF STRATEGY, THE DESIGN AND TESTING OF COMPUTER PROGRAMS, AND IN OTHER FIELDS WHERE CALCULATION WITH NUMBERS IS LESS IMPORTANT THAN CALCULATION WITH SYMBOLS. NOT ONLY IS LISP A LANGUAGE FOR PROGRAMMING COMPUTERS BUT IT IS ALSO A MATHEMATICAL-SYSTEM 3, IN THE SAME WAY AS THE GEOMETRY OF EUCLID 1 OR THE ORDINARY ALGEBRA OF NUMBERS IS A MATHEMATICAL-SYSTEM.

LISP DIFFERS FROM MOST PROGRAMMING LANGUAGES IN SEVERAL IMPORTANT WAYS. ONE DIFFERENCE IS THAT ALL DATA 3 THAT LISP DEALS WITH CONSISTS OF SYMBOLIC-EXPRESSIONS 3. THE SYMBOLIC-EXPRESSIONS THAT ARE ACCEPTABLE TO LISP ARE CALLED "S-EXPRESSIONS 3".

A SECOND DIFFERENCE IS THIS: WHEN PROGRAMMING IN LISP, WE DO NOT HAVE TO ALLOCATE PORTIONS OF THE MEMORY OF THE COMPUTER FOR STORING SYMBOLIC-EXPRESSIONS; THIS IS DONE AUTOMATICALLY BY THE IMPLEMENTATION 3\ OF LISP ON THE COMPUTER.

A THIRD IMPORTANT DIFFERENCE IS THAT LISP CAN INTERPRET 3 AND EXECUTE PROGRAMS WRITTEN IN LISP. THUS, LIKE MACHINE-LANGUAGE 3 AND UNLIKE MANY OTHER PROGRAMMING LANGUAGES, LISP CAN BE USED TO GENERATE 3 AND EXECUTE 3 NEW PROGRAMS. IN ADDITION, THE WAY IN WHICH LISP INTERPRETS 3 AND EXECUTES 3 LISP CAN BE WRITTEN IN LISP.

AS WE SAID ABOVE, SYMBOLIC-EXPRESSIONS ACCEPTABLE TO LISP ARE CALLED S-EXPRESSIONS. (THE LETTER -S- IN S-EXPRESSION 3 COMES FROM THE FIRST LETTER OF THE WORD SYMBOLIC.)
WHAT MAKES THEM ACCEPTABLE?

THE SIMPLEST KIND OF SYMBOLIC-EXPRESSION 3 ACCEPTABLE TO LISP IS ONE WHICH REGULARLY DOES NOT HAVE ANY PARTS RECOGNIZABLE BY LISP AND IS TREATED BY LISP AS A UNIT. AN ACCEPTABLE SYMBOLIC-EXPRESSION OF THIS KIND IS CALLED AN ATOMIC-SYMBOL 3.

IN ALMOST ALL VERSIONS OF LISP, CAPITAL LETTERS WRITTEN TOGETHER WITHOUT SPACES (IN THE SAME WAY AS LETTERS ARE WRITTEN TOGETHER TO MAKE WORDS IN ORDINARY ENGLISH) ARE TREATED AS ATOMIC-SYMBOLS 3 ; FOR EXAMPLE, PRINT 4, PLUS 4, NULL 4. ALSO CAPITAL LETTERS & DIGITS WRITTEN TOGETHER IN A GROUP IN WHICH THE FIRST CHARACTER IS A CAPITAL LETTER IS AN ATOMIC-SYMBOL: FOR EXAMPLE PRIN1 4.
IN MANY VERSIONS OF LISP, NUMBERS ARE TREATED AS ATOMIC-SYMBOLS 3 OF A SPECIAL KIND, SINCE THEY HAVE SOME PROPERTIES THAT MOST ATOMIC-SYMBOLS DO NOT HAVE: FOR EXAMPLE, 22, 395. THERE IS A UNIQUE ATOMIC-SYMBOL 3 REPRESENTED AS NIL OR (), A PAIR OF PARENTHESES WITH NOTHING 2 INSIDE. IT STANDS FOR AN EMPTY LIST, A LIST WITH NO ELEMENTS. EXAMPLES OF ATOMIC-SYMBOLS ARE: PLUS 4, 22 1, M4 4, NIL 3, TRANSITIVERELATION 4.

ALL S-EXPRESSIONS 3 ARE EITHER ATOMIC-SYMBOLS OR THEY ARE BUILT UP OUT OF ATOMIC-SYMBOLS IN SPECIFIED WAYS.

ONE WAY IS THIS: PUT TOGETHER A LEFT PARENTHESIS, THEN AN ATOMIC-SYMBOL OR S-EXPRESSION, THEN A DOT, THEN AN ATOMIC-SYMBOL OR S-EXPRESSION, THEN A RIGHT PARENTHESIS, AND THE RESULT IS AN S-EXPRESSION. THIS KIND OF S-EXPRESSION 3 USES WHAT IS CALLED DOT-NOTATION 3. THE DOT 3 STANDS FOR AN OPERATOR OF LISP 3 WHICH IS CALLED CONS 4 (PRONOUNCED "CONSS" 4). EXAMPLES OF S-EXPRESSIONS IN DOT-NOTATION 3 ARE: ..... SINCE "S-EXPRESSION" INCLUDES "ATOMIC-SYMBOL", "AN ATOMIC-SYMBOL OR S-EXPRESSION" IS THE SAME AS "AN S-EXPRESSION".

A SECOND WAY IS THIS: PUT TOGETHER A LEFT PARENTHESIS, AN S-EXPRESSION, A SPACE, AN S-EXPRESSION, A SPACE, AN S-EXPRESSION, ... ,A SPACE, AN S-EXPRESSION, AND A RIGHT PARENTHESIS, AND THE RESULT IS AN S-EXPRESSION. THIS KIND OF S-EXPRESSION USES WHAT IS CALLED LIST-NOTATION 3. ANY SUCH EXPRESSION IS EQUIVALENT TO AN EXPRESSION WRITTEN IN DOT-NOTATION ACCORDING TO CERTAIN RULES WHICH WILL BE EXPLAINED LATER.

EXAMPLES OF S-EXPRESSIONS IN LIST-NOTATION ARE:

.....

ALL S-EXPRESSIONS CAN BE WRITTEN IN DOT-NOTATION. NOT ALL S-EXPRESSIONS CAN BE WRITTEN IN LIST-NOTATION.

Exhibit 10 A

THE SPECIAL VOCABULARY OF THE REVISED EXPLANATION

AS OUTPUT BY THE COMPUTER PROGRAM

```
CLASS 3
ATOMIC-SYMBOL 7
ATOMIC-SYMBOLS 6
DATA          1
DOT           2
DOT-NOTATION 4
EXECUTE       2
EXECUTES      1
GENERATE      1
IMPLEMENTATION 1
IMPLEMENTED 1
INTERPRET     1
INTERPRETS 1
LISP          26
LIST-NOTATION 3
MACHINE-LANGUAGE 2
MATHEMATICAL-SYSTEM 2
NIL           2
S-EXPRESSION 14
S-EXPRESSIONS 7
SYMBOLIC-EXPRESSION 2
SYMBOLIC-EXPRESSIONS 6
NU 21 FQ 92
```

Exhibit 10 B

THE SPECIAL VOCABULARY OF

THE REVISED EXPLANATION

POST-EDITED, SUMMARIZED, AND CLASSIFIED

Word	Frequency	Explanation Subclass
atomic symbol(s)	13	B
data	1	D
dot	2	D
dot notation	4	D
execute (a program)	3	P
generate (programs)	1	P
implement(ed)	1	P
implementation	1	P
interpret (programs)	2	P
LISP	26	B
list notation	3	D
machine language	2	P
mathematical system	2	D
NIL	2	D
S-expression(s)	21	B
symbolic expression(s)	8	D
( )	2	D

Basic term
Probably understandable
Defined term

Number, 17 Frequency, 94 Summary, B, 3
D, 7
P, 7

Exhibit 11 A

THE ALLUSTON VOCABULARY OF THE REVISED EXPLANATION

AS OUTPUT BY THE COMPUTER PROGRAM

```
CLASS 4
CONS          1
CONSS         1
DIFFERENTIAL-AND-INTEGRAL-CALCULUS 1
IBM-7090      1
M4            1
MATHEMATICAL-LOGIC 1
NULL          1
PLUS          2
PRIN1         1
PRINT         1
THEORY-OF-ELECTRICAL-CIRCUITS 1
TRANSITIVERELATION 1
NU 12 FQ 13
```

Exhibit 11 B

THE ALLUSION VOCABULARY FOR THE REVISED EXPLANATION

POST-EDITED AND SUMMARIZED

Word	Frequency
differential and integral calculus	1
theory of electrical circuits	1
mathematical logic	1
PRINT	1
PLUS	2
IBM 7090	1
NULL	1
PRIN1	1
M4	1
TRANSITIVERELATION	1
CONS	1
conss	1
(A . 22)	1
(B . (A . C))	1
(2 . (4 . (6 . ( ))))	1
((2 . 22) . ((3 . 33) . ((4 . 44))))	1
((A . B) . C)	1
(A . NIL)	1
(A . (B . NIL))	1
(A B C D E)	1
(B)	1
(PLUS 22 33 55)	1
(REVERSEOF (S U L P))	1
((W (K 50) (P 51)) (B (K 43) (P 54))))	1
((2 3 6) (2 4 8) (3 5 15) (4 7 28))	1

Number 24 Frequency 25

Exhibit 12 A

THE KEY VOCABULARY OF THE REVISED EXPLANATION

AS OUTPUT BY THE COMPUTER PROGRAM

CLASS	2
ABOVE	1
ACCEPTABLE	5
ACCORDING	1
ADDITION	1
ALGEBRA	1
ALLOCATE	1
ALMOST	1
ALSO	2
AMONG	1
ANY	2
AUTOMATICALLY	1
CALCULATION	2
CALCULATIONS	1
CAPITAL	3
CERTAIN	1
CHARACTER	1
CLASSIFIED	2
COMPUTER	6
COMPUTERS	2
CONSISTS	1
DEFINED	1
DESIGN	1
DIFFERENCE	3
DIFFERS	1
DIGITS	1
EITHER	1
ELEMENTS	1
EMPTY	1
ENGLISH	1
EQUIVALENT	1
EXAMPLE	4
EXAMPLES	3
EXPLAINED	1
EXPRESS	1
EXPRESSION	2
GEOMETRY	1
IDEAS	1
IMPORTANT	3
INCLUDED	1
INCLUDES	1
INSIDE	1
LANGUAGE	1
LANGUAGES	3
LATER	1
LETTER	4
LETTERS	4
MANY	2
MEMORY	1
NOTHING	1
NUMBERS	7
OFTEN	1
ONLY	1
OPERATIONS	1
OPERATOR	1
ORDINARY	2
OTHER	4
OTHERS	1
PARENTHESES	1
PARENTHESIS	4
PARTICULARLY	1
PORTIONS	1
PROCESS	1
PROCESSING	1
PROGRAMMING	5
PROGRAMS	3
PRONOUNCED	1
PROPERTIES	1
PURPOSE	1
RAILROAD	1
REASON	1
RECOGNIZABLE	1
REGULARLY	1
REMARKABLE	1
REPRESENTED	1
RESULT	2
SCIENTISTS	1
SECOND	3
SEQUENCE	1
SEVERAL	1
SIMPLEST	1
SPECIAL	1
SPECIFIED	1
STATIONS	2
STRATEGY	1
SYMBOLIC	1
SYMBOLS	6
TABLES	1
TIMETABLE	1
TOGETHER	6
UNIQUE	1
UNIT	1
UNLIKE	1
USEFUL	1
VERSIONS	2
WITHOUT	1
WRITTEN	8
NU 96 FQ	172

Exhibit 12 B

THE KEY VOCABULARY FOR THE REVISED EXPLANATION

POST-EDITED AND SUMMARIZED

Word	Frequency	Word	Frequency
above	1	letter(s)	8
acceptable	5	many	2
according	1	meaning	1
addition	1	memory	1
algebra	1	nothing	1
allocate	1	number(s)	1
almost	1	often	1
also	2	only	1
among	1	operation(s)	1
any	2	operator	1
automatically	1	ordinary	2
calculation(s)	3	other	3
capital	3	others	1
certain	1	parentheses	1
character	1	parenthesis	4
classified	2	particularly	1
computer(s)	8	portion(s)	1
consist(s)	1	process(ing)	2
defined	1	program(s)	3
design	1	programming	5
differ(s)	1	pronounced	1
difference	3	properties	1
digit(s)	1	purpose	1
either	1	railroad	1
element(s)	1	reason	1
empty	1	recognizable	1
English	1	regularly	1
equivalent	1	remarkable	1
example(s)	7	represented	1
explained	1	result	2
express	1	scientist(s)	1
expression	2	second	3
geometry	1	sequence	1
idea(s)	1	several	1
important	3	simplest	1
include(s)	1	special	1
inclu[illegible]	1	specified	1
inside	1	station(s)	2
language(s)	4	strategy	1
later	1	symbol(s)	6

Word	Frequency
symbolic	1
table(s)	1
timetable	1
together	6
unique	1
unit	1
unlike	1
useful	1
version(s)	2
without	1
written	8
....	1

Number, 92; Frequency, 171

None of these words can readily be eliminated by a shorter or simpler word.

Exhibit 13

THE ONE-SYLLABLE VOCABULARY OF THE REVISED EXPLANATION

CLASS	1	LEAVE	1	TESTING	1
-S-	1	LEFT	2	THAN	1
1958-60	1	LESS	1	THAT	10
22	2	LIKE	1	THE	34
395	1	LIST	5	THEM	1
A	27	LISTS	7	THEN	4
ALL	5	M-L-MINSKY	1	THERE	1
AN	17	MAKE	1	THEY	4
AND	16	MAKES	1	THIRD	1
ARE	14	MAY	2	THIS	7
AS	8	MEANING	1	THREE	1
AT	1	MIT-CAMBRIDGE-MASS	1	THUS	1
BE	8	MORE	2	TIMES	1
BEEN	1	MOST	2	TO	9
BUILT	1	NAME	2	TRAIN	2
BUT	3	NAMES	2	TRAINS	1
BY	6	NEW	2	TREATED	3
CALLED	8	NO	1	TWO	1
CAN	5	NOT	6	UP	1
CHIEF	1	OF	33	USED	3
COMES	2	ON	2	USES	2
COULD	1	ONE	5	WAS	1
DEALS	1	OR	11	WAY	7
DO	2	OUT	2	WAYS	3
DOES	1	PAIR	1	WE	2
DONE	1	PARTS	1	WHAT	3
EUCLID	1	PLAYING	1	WHEN	1
FIELDS	1	PUT	3	WHERE	1
FIRST	4	RIGHT	2	WHICH	8
FOR	13	RULES	1	WILL	1
FROM	3	RUN	1	WITH	7
GAMES	1	S-B-RUSSELL	1	WORD	1
GROUP	3	SAID	1	WORDS	2
HAS	1	SAME	2	WORKED	1
HAVE	4	SINCE	2	WORKING	1
HAVING	1	SO	1	WOULD	1
IN	28	SOME	2	NU 119 FQ 485	
IS	30	SPACE	3		
IT	3	SPACES	1		
JOHN-MCCARTHY	1	STANDS	2		
KIND	5	STORING	1		
KINDS	1	SUCH	1		

Exhibit 14

COMMON PROPERTIES AND RELATIONS OF IDEAS -- CHECKLIST

Name, identification
Other names, repetition in other words, equivalents
Examples, instances
Definition, meaning, significance
Essence, theme, nature

Kind, sort, genus, species, class
Properties, nature, habits
Similar things, related things, associated things
Opposites, contrasts
Distinguishing characteristics

Things included in it, parts
Things of which it is a part
Context, environment, situation, field
Composition, material, substance
Structure, organization, construction

Activity, behavior, verb
Agents, doers, subject of verb
Products, object of verb, recipients
Manner, ways, adverbs
Size, dimensions, measurements

Quantity, number
Variation, range, average, deviations
Shape, form, solid liquid or gas
Weight, density
Appearance, look, color, luster

Sound, smell, taste, feel
Place, location, position, extent, prevalence
Time, duration, age, persistency
History, origin, causes, development
Future, results, effects, predictions

Purpose, function, use, worth, value
Advantages, disadvantages
Owners, users
Importance, relation to human affairs

Chapter 8

Judging and Improving Two More Sample Explanations

Outline

1. Assessing "Principles of Sampling"
2. Revising the Explanation
3. Judging the Revised Explanation
4. The Audience for "Screening out Potential Troublemakers"
5. Assessing the Explanation
6. Revising the Explanation
7. Judging the Revised Explanation

So far, we have taken one sample explanation and shown -- at very great length -- how it may be revised and improved.

Let us now take two more sample explanations, consider how they can be revised and improved, and concentrate on describing the revising process. This time we will place the original and revised explanations and their vocabulary classifications in Appendix 2.

The first of these two additional sample explanations deals with principles of sampling, and is Explanation 4 in Appendix 2. The original explanation, "Principles of Sampling", is taken from a report written by three distinguished statisticians, and published by the American Statistical Association. The report was originally prepared for the National Research Council, but was subsequently published and made available to a wide audience of interested persons, including particularly scientists. The revised explanation is stated as Explanation 5 in Appendix 2.

1. Assessing "Principles of Sampling"

Examining the output of the computer program and doing a little arithmetic we obtain the following summary:

Vocabulary	Class	Terms		Percentages	
Name	No.	Number	Frequency	Number	Frequency
Special terms	3	29	50	11.4%	10.2%
Allusions	4	0	0	0	0
Key vocabulary	2	110	138	46.2	28.2
One-syllable vocabulary	1	108	301	42.4	61.6
	Total,	255	489	100.0	100.0

Only a few of the words in Class 2 are disturbing: "carapaces," "semblance," "tenuous," each used only once. The word "carapaces" occurs in a context where it does not have to be understood -- we could have marked it Class 4 if we had wished. The word "tenuous" will need attention.

Let us take a look at the 29 special terms put into Class 3. The numeral following the slant sign (/) reports the frequency of occurrence of the term in the sample explanation.

The following special terms can probably be understood (or correctly guessed) by the audience:

drawn / 2

on the average / 1

sample(s), sampling, sampling plans (probably "plans for sampling") / 8

The following special terms occur in the explanation with one or more statements (or phrases) which define them (even if they may be defined in terms of undefined terms):

grab sample(s) / 5

mechanical randomization / 1

non-random samples / 1

The following terms need to be either defined or eliminated because in statistics they have precise technical meanings, which the audience will not necessarily know and may not guess correctly:

random, randomly, at random / 4

random number tables (Are they "number tables" which are "random"?) / 1

random samples (Is a "random sample" also a "simple random sample"?) / 2

simple random sample(s) / 5

bias / 1

population(s) / 3

existent population (When is a population "existent"?) / 1

independent / 1

mean(s) / 5

probable error (Is it an "error" which is "probable"? / 1

range / 1

valid / 1

variability / 3

fluctuation (Is it the same as "variability"?) / 2

stability (Is it the opposite of "variability"?) / 1

It could be argued that much of the audience would absorb at least some idea from most of these terms. But to produce a really good explanation, most of these terms should be made completely clear.

Because of the frequency of the occurrence of the ideas, it is clear that the following ideas are basic to the discussion:

sampling / 8

randomness / 13

population(s) / 4

mean(s) / 5

variability / 6

Among the terms which occur only once or twice, the following are candidates for elimination:

mechanical randomization / 1

existent population / 1

fluctuation / 2

stability / 1

2. Revising the Explanation

In revising any explanation, we do not want to change the flavor of the explanation; where colorful expressions are used, we want to keep them. We desire only to increase the clarity of the ideas referred to, so that the audience may clearly understand what is being said.

1. "by biologists, sociologists, engineers, or chemists": 16 syllables, 5 words. This phrase is a mouthful. Furthermore, the class of scientists mentioned is not complete -- many more kinds of scientists are implied. Let us change the phrase to "In scientific investigation".

2. "In the early years of the present century": 12 syllables, 8 words. The idea is hardly worth this much. Let us change it to "Fifty years ago", 4 syllables, 3 words.

3. "carapaces": 4 syllables. This is a rare word. The dictionary definition mentions "shell". There is a pleasant sound to "claws and carapaces" but this alliteration does not really belong in scientific writing of this kind. Let us use "shells", one syllable, one word.

4. "results": What results? The results might be measurements. But 1000 measurements would not each have a figure attached to them. From knowledge of statistics, we know that most likely the "results" would be the average or the arithmetical mean of the measurements -- the sum of all the measurements divided by the number of them -- also called simply "the mean".

5. "probable error": This term is in quotes. This tells the reading audience (a listening audience cannot hear quotes) that the term has a special meaning. But no clue is given to the special meaning. What does the author mean?

In statistics, the "probable error" is defined as equal to .6745 times the "standard deviation". What is "standard deviation"? The standard deviation is the square root of the sum of the squares of the deviations of measurements from the arithmetic mean divided by the number of measurements. As a statistical figure, however, the "probable error" is very old-fashioned; what is regularly used nowadays for several good reasons is the standard deviation.

The "probable error" received its name from the fact that under many conditions of sampling, it is equally probable that a measurement would fall inside one probable error from the mean as it would fall outside. But nowadays the corresponding idea is used as follows: Under many conditions of sampling, the chance is 19/20 that a deviation will fall within 2 standard deviations from the mean and 997/1000 that it will fall within 3 standard deviations from the mean.

To put all this information into the explanation is not desirable. We don't have the space, and the idea is not worth that much attention. What shall we do?

Let's use a phrase to identify the kind of thing that "probable error" is. A reasonable wording is:

a computed degree of variation called the "probable error".

6. Long sentence. The sentence in which "probable error" occurs contains 64 words. We will do well to break it. A reasonable place to break is at the "which". We can start a new sentence with "This figure". The new sentence now reads: "This figure could have been appropriate had the 1000 crabs or the 1000 leaves been drawn at random from the population of interest."

7. "had": This "had" in inverted order signals a condition. The common everyday form of a condition uses "if". Let's use "if".

8. "drawn at random": This is a technical phrase referring to a well-defined

operation in statistics, which we need to make very clear to the audience. When we sample, we take or select or "draw" instances, examples, cases. But what does "at random" mean?

"Webster's New Collegiate Dictionary" published by G. and C. Merriam Co., Springfield, Mass., 1961, 1174 pages, gives on page 700:

> random noun. (From the Old French, randon, violence, rapidity.) Now Rare. A haphazard course or progress. -- at random. Without definite aim, direction, rule, or method; at haphazard. -- adjective. 1. Coming, acting, made, occurring, etc., at random. .... Syn. Random, haphazard, casual, desultory mean showing the influence of accident rather than design. Random implies little or no guidence by a governing mind, eye, objective, or the like; haphazard, a being more or less at the mercy of chance or of natural or logical necessity; casual, a working, an acting, or the like, without deliberation, intention, or purpose; desultory, a jumping or skipping from one thing to another, ungoverned by method or system.

None of these definitions states clearly what the statistician means by "randomness". For example, on page 100 of "Probability and Statistics" by F. Mosteller, R. E. K. Rourke, and G. B. Thomas, Jr., Addison Wesley Publishing Co., 1961, appears:

> The fair tossing, the thorough shuffling, and the blindfold drawing are physical processes that we use in trying to achieve what is called "randomness"; that is, trying to give all outcomes equal chances ...

We must emphasize this idea since it is different from the ordinary meaning of "random".

So we can say:

> drawn at random, that is, under conditions where each individual has an equal chance of being drawn

What is "an equal chance"? The assumed audience consists of persons with scientific training. People are familiar with "a chance" from childhood. It is reasonable to expect that "an equal chance" will be understandable to this audience.

9. population of interest. The word population regularly makes us think of people, often the people of a country or a region. The phrase "of interest" means "that we are interested in", but it has an academic flavor. But what does a statistician mean by "population"?

The dictionary gives:

> population noun. .... 4. Biology. The organisms, collectively, inhabiting an area or region. 5. Statistics. A group of individuals or items; specifically, in biometry, the entire group of organisms from which samples are taken for measurement.

So here again we need to put the audience on the right track to understand the technical meaning:

> the group of all individuals in which we are interested, which we are studying, and which we are drawing samples from

We can shorten that in this context to:

> the group of all individuals being studied

The sentence now becomes:

> This figure could have been appropriate if the 1000 crabs (or leaves) had been drawn (or selected) from their "population" (the group of all individuals being studied), "at random", that is, under conditions where each individual has an equal chance of being drawn.

The sentence is much longer, but it contains three "reminder" definitions as keys to help the audience.

10. such actions. What actions? The actions of measuring? No. The actions of attaching probable errors? No. The actions of not drawing at random -- but drawing individuals as they happened to come to hand, and without considering

whether each individual in the population had an equal chance to get into the sample. How shall we make this ambiguous reference unmistakably clear? Perhaps we can do it with an extra sentence -- more words, unfortunately, but in a good cause:

> But nearly all the time the condition of drawing at random was not observed.

Now we can start the next sentence with "Such drawings", and we have a clear reference:

> Such drawings were unwarranted shotgun marriages between the quantitatively unsophisticated idea of a sample as "what you get by grabbing a handful" and the mathematically precise notation of a "simple random sample".

The colorful language (down to the last three words) can be kept exactly as is, because it will certainly be clear to the audience.

11. "simple random sample". Here again quotes are used, to show that the phrase is a special term. We receive one more piece of information about it, that it is a mathematically precise notion. But we are not told the notion, and the audience should be told. Let's say what it is, and decrease the mystery:

> a "simple random sample", a sample where each individual in the population of interest has an equal chance of being drawn".

We put back in the academic phrase "of interest" to help preserve the flavor of the authors' writing, and because now the technical meaning of the word population has been made clear previously.

12. In the years between. Now we have to pay a penalty for changing "in the early years of the present century". We said above "fifty years ago" and now there are no "years between". Let's say "since that time".

13. semblance. I think this is a wrong word. The dictionary definition of "semblance" is:

> semblance noun. (Old French) 1. One's outward appearance; form. 2. Countenance; aspect. 3. An image; likeness. 4. Resemblance, actual

or apparent; similarity. 5. Seeming; especially, specious appearance or seeming; also, mere show.

Clearly a "semblance" is not enough. We have to have "some real steps in the direction of". Let's say "approach".

14. mechanical (dice, coins, random number tables, etc.) randomization. This phrase is a mouthful. It is reasonable to rewrite it, making the ideas clear. We put:

> We insist on at least some approach towards using a mechanical device for selecting at random, a device such as coins or dice or tables of numbers that occur at random, in order to draw individuals for the sample.

15. before. The sentence has become too long, and we break it here, making a new sentence starting with the words "only then".

16. existent population. Here the authors have devoted just one word, "existent", to an idea of much significance: that there are some populations that do not actually exist where any sample you draw is a random sample. For example, when you toss a coin fairly 50 times, the sequence of your 50 outcomes heads or tails is a perfectly satisfactory simple random sample of the infinite population of all possible fair tossings of fair coins, and you have no problem of avoiding samples that are not random. On the other hand when a population does exist, it exists in space and time, and so does the investigator, and some of the individuals of the population are much more accessible to him than others, and so he has to design his sampling to avoid the bias accessibility may give.

It would be a help to the audience to devote more words to this important idea. We can say:

> a population that exists (some populations are indefinitely large and do not exist)

17. members. What are "members"? In the early part of this sentence the

word "individuals" was used for this idea. Let's stick to the word "individuals": the same word for the same idea is often helpful in understanding.

18. entering the sample. What is "entering"? Clearly, the individuals of a statistical sample may not have power to act on their own. What is intended is "being drawn for". Let's say that.

19. Since ..., it follows .... Here is a long sentence with many ideas to be understood. We will be better off if we break it and say "...; so it follows ...."

20. means. This word refers to arithmetical averages, as mentioned above. Let us put in a "reminder" definition:

the means* (the averages)

to help the audience stay on the track of the idea which the authors intend.

21. of the same size. The same size as what? There is an implied condition and comparison. So what is being asserted could well be stated a little more carefully:

the means of grab-samples of a given size resemble one another less than the means of random samples of that size.

22. variability. What is "variability"? The property of being variable? No. The degree to which something is variable? No. This is a statistical idea, and the authors are thinking of the measurement of the amount of variation or spread in a set of statistical observations. Let's select one phrase and stick to it, wherever this idea is referred to. Suppose we use "degree of variation", which for this audience may be a little more understandable than "degree of variability".

23. although .... stability. These two sentences express some rather complicated ideas statistically, which can hardly be adequately understood without some professional knowledge of statistics. So let us omit some portions of the idea and say only the rest:

Applying to grab-samples the formulas that are appropriate only for random samples and putting into those formulas the means of grab-samples,

introduces bias. This leads to an unwarranted appearance of a smaller degree of variation in the population than actually exists.

Instead of "stability" we have used "smaller degree of variation". Instead of double bias, we have referred only to bias.

24. bias. What is "bias"? From the dictionary, we have:

bias noun. 2.c. Statistics. A tendency of an estimate to deviate in one direction from a true estimate (as by reason of non-random sampling).

A short reminder definition of this technical idea, which seems sensible to use in this place, is:

bias (error in one direction)

25. independent. What are "independent" chances? Probably not all of the audience knows what "independent chances" are. On the other hand, a fair amount of the notion can be gathered from the ordinary meaning of "independent", that is, not depending on something else. Although it would be desirable to explain this idea to the audience and nail it down, this would probably go too far afield. So regretfully we pass by the word without clarifying it.

26. likely range of fluctuation. The authors are referring to variability which we are calling "degree of variation." So let's say:

the likely degree of variation

27. sampling plans. Often a noun is put as an adjective in front of another noun, and the preposition which previously connected them is suppressed. On these occasions there is always an increase of possible confusion. The phrase "sampling plans" sounds as if it might have a technical meaning. Let us say therefore "plans for sampling".

28. measure of fluctuation. Let's translate this into "degree of variation."

29. non-random samples. Let us say "samples that are not random".

30. entirely tenuous. "Tenuous" is a rare word. Also the meaning of this sentence is obscure unless the meaning of "tenuous" is known. The sentence is like

"she was wearing a wimple". The dictionary gives for the meanings of "tenuous":

1. Thin; slender. 2. Rare; subtile; not dense. 3. Unsubstantial; insignificant; flimsy."

Let's say:

without a shred of justification

31. principles involved. What "principles"? Clearly, the principles of sampling. Let's say so.

3. Judging the Revised Explanation

Again, most of the work of summarizing the vocabulary can be accomplished by the computer program.

The summary of the output appears in the following table:

Vocabulary Class		Terms		Percentage	
Name	No.	Number	Frequency	Number	Frequency
Special terms	3	23	61	8.4	10.6
Allusions	4	0	0	0	0
Key vocabulary	2	125	156	47.7	27.1
One-syllable vocabulary	1	115	359	43.9	62.3
	Total	262	576	100.0	100.0

The special terms (grouped a little) are:

random, at random, randomly / 7

random samples / 4

simple random sample / 3

averages, on the average / 3

bias / 1

degree of variation / 5

draw, drawing, drawings, drawn / 8

grab samples / 5

independent / 1

means / 3

population(s) / 7

probable error / 1

sample(s), sampling / 13

Every one of these terms is accompanied by one or more defining or helping statements or phrases.

There are no words in the key vocabulary (class 2) which might disturb the audience.

It seems evident that the explanation is improved.

4. The Audience for "Screening out Potential Troublemakers"

The second of the two sample explanations is Explanation 6 in Appendix 2. It is taken from pages 40 and 41 of Chapter 8, "The Case of the Suspended Sentence", in "Principles and Problems of Naval Leadership", second edition, NAVPERS 15924A, 1964, 107 pages, prepared by the Bureau of Naval Personnel, available from the Superintendent of Documents, Washington, D.C., 70 cents. This publication:

> attempts to synthesize some of the ideas developed in the sciences of psychology, human relations, management, morals, and ethics and present them in a Navy context. .... The principles of effective naval leadership are demonstrated in this text through the case-study method.

The problem discussed in this chapter is what to do in a case where a certain sailor violated a suspended sentence period by going absent without leave, and upon return stated that he desired a bad conduct discharge.

The discussion is presented as answers to three questions:

1. Why doesn't the Navy screen prospective recruits to keep the troublemakers out of the Navy?
2. At what point in a man's "disciplinary career" should steps be taken

to prevent further offenses?

3. When should a man be separated from the naval service for his own good as well as the good of the Navy?

The corresponding headings of the discussion are:

1. Screening out Potential Troublemakers
2. The Common Pattern of Typical Troublemakers
3. Effects of Certain Types of Disciplinary Action

Here we shall restrict ourselves to the sample explanation consisting of the whole discussion under the first heading only.

The audience for this explanation is made clear from some of the statements in the introduction to the text:

> The cases and problems are actual ones submitted by experienced officers. ... The principles treated here are considered to be simple, realistic, and practical. Properly understood, they can serve as a basis of practical operation for the naval officer charged with the responsibilities of leadership, as this term is understood in the Naval Establishment.

The way in which the explanation is to be read and understood is further expressed by some more statements in the introduction:

> All the cases taken together produce a body of principles and a pattern of approach to most leadership problems. The technical terms of psychology, sociology and scientific management have been avoided. ....The concepts underlying these terminologies .... have been to the greatest extent feasible translated into the working language of the average naval officer.

In considering this explanation therefore, we can conclude that the audience for which it is prepared consists of Naval line officers (in contrast to Naval staff officers) who are concerned with leadership and with the supervision and management of men.

5. Assessing the Explanation

Examining the output of the computer program and doing a little arithmetic we obtain the following summary:

Vocabulary Class		Terms		Percentages	
Name	No.	Number	Frequency	Number	Frequency
Special terms	3	16	37	4.2	4.5
Allusions	4	9	9	2.4	1.1
Key vocabulary	2	181	235	47.6	28.7
One-syllable vocabulary	1	174	530	45.8	65.7
Total		380	819	100.0	100.0

This is a low percentage of special terms and allusions; and we have no criticism at this point.

We now examine the vocabulary used in the original explanation in the following way:

-- list the 16 terms appearing in Class 3, and comment on them,

-- list the 9 terms appearing in Class 4, and comment on them.

-- consider the terms appearing in Class 2; select those which may give some difficulty (four);, and comment on them.

This is done in Exhibit 1 at the end of this chapter.

Summarizing Exhibit 1, we determine that the following 13 terms are not used well in the explanation; they should be either eliminated or changed, or their use should be changed:

computers

cybernetics

we, our, us

personal moral responsibility

chi-squares

coefficients of correlation

electroencephalograms

psychogalvanometers

Rorschach diagrams

fluoroscopes

damnation

exhausting

full-blooded

6. Revising the Explanation

We now set to work to revise and rewrite the explanation.

In the original explanation, the main propositions stated or implied are shown below:

No.	Main Proposition	Paragraph
1	The behavior of human beings is far from completely determined by scientific or other laws.	1, 2, 4
2	The behavior of individual human beings is very variable.	2, 3, 4
3	Even if one could predict that recruits coming, say, from broken homes would become troublemakers, nevertheless some of the Navy's most successful personnel come from such homes.	5
4	Tests that seek to identify potential troublemakers have not been successful.	6
5	Even if there were successful predictive tests, it might be difficult for the Navy to inform some applicants that their test scores were insufficient. (See Note 1)	7
6	Because manpower supply is limited, the Navy has to accept some persons with poor backgrounds.	8
7	In periods of mobilization, even more persons with poor backgrounds have to be accepted.	8
8	Screening out potential troublemakers is important because of the high cost of disciplinary problems.	9

9	Once the Navy has poor recruits, it should do the best it can with them. (See Note 2)	9
10	Some predictive factors bearing on troublemaking have been observed: coming from a broken home, a record of arrest, failure to complete high school.	5, 8

Note 1. Proportion of Words. It is clear that too many words (four paragraphs) are devoted to propositions 1 and 2. Furthermore, the discussion here is the source of all the words in Vocabulary Class 4, which though colorful are not necessary and may be confusing.

Note 2. Invalid Argument. Proposition 5, that it might be difficult for the Navy to inform applicants that their test scores were too low for entry, is not a valid argument. For over and over again the Navy rejects applicants for low scores on tests, and the Navy is not required to state all the implications of low scores.

Note 3. Irrelevant Statement. Proposition 9, that the Navy should do what it can with poor recruits, is not really relevant to the subject "screening out potential troublemakers". But it is a reasonable proposition to include, since the author is pointing out that it is both difficult and impractical to screen out all potential troublemakers.

From the foregoing, it is clear that the original explanation has a number of structural defects: (1) Propositions 1 and 2 are covered too voluminously; (2) the propositions are not arranged in a logical sequence. Let us try to arrange the propositions in a more logical sequence.

Old No.	Rearranged Sequence of Main Propositions (with Connections)	New No.
8	Screening out potential troublemakers is important because of the high cost of disciplinary problems. BUT	1
4	Tests that seek to identify potential troublemakers have not been successful. BECAUSE	2
1	The behavior of human beings is far from completely determined by	3

scientific or other laws. AND

2 The behavior of individual human beings is very variable. ---- 4

10 Some predictive factors bearing on troublemaking have been observed, 5
such as coming from a broken home, a record of arrest, or failure
to complete high school. BUT EVEN SO

3 Some of the Navy's most successful personnel come from broken homes. 6

AND

5 Because manpower supply is limited, the Navy has to choose persons with 7
poor backgrounds. AND

7 In periods of mobilization even more persons with poor backgrounds have 8
to be accepted. SO

9 Once the Navy has poor recruits, it should do the best it can with them. 9

Other points that have been attended to in the process of revision are as follows:

Narrowing of Reference of Pronouns. The word "we, our, us" should be restricted to mean precisely "Navy officers", which is the natural and desirable meaning in this situation. Each other occurrence of the word should be replaced by naming the group of persons intended.

Removal of Exaggeration. A number of sentences should be revised from partially true (or true only if interpreted in a restricted way) into exactly true sentences.

Example 1: Original: Many of those who don't look promising can and have become the finest sailors in the fleet.

Revised: Many of those who did not look promising in the past are now among the finest sailors of the fleet.

Comment: The revision by the use of the word "among" allows some of "the finest sailors in the fleet" to come from those who looked promising at the start.

Example 2: Original: He can obey the law for 50 years, then murder someone.

Comment: 50 years is a long time in a man's life, say age 17 to 67, and would probably be outside of observation of men on active duty in the Navy.

Revision: He can obey the law for 30 years, then murder someone.

Example 3: Original: We know of none ((good tests)) to date, and efforts to develop one have been unsuccessful.

Revision: Tests ... to date have not been successful in this sort of prediction.

Trimming. By trimming we mean increasing the presision of words and statements.

Many unnecessary words can be omitted (provided the substance of the explanation is retained) such as references to advanced or unclear ideas, including cybernetics, refined computers, psychogalvanometers, Rorschach tests, the primrose path to damnation, etc.

Many vague or sweeping statements can also be omitted:

People remain people. (What does that mean?)

A man can know all the rules. (What is the implication?)

No machine can have a built-in personal moral responsibility. (Neither do persons have one that is built-in. Besides, who can prove what will happen in the programming of machines over even the next 200 years?)

We are now ready to rewrite the explanation. We have two guides:

(1) what needs to be said to smooth out the presentation of the propositions in the revised, more logical sequence; and

(2) the original explanation, which will suggest ideas and wording

For we need of course to take from the original explanation those ideas and ways of speaking that are important and clear, so that we preserve the flavor of what is here being said.

The result of the rewriting appears as Explanation 7 in Appendix 2.

7. Judging the Revised Explanation

We are now ready to assess the revised explanation.

Examining the output of the computer program, and doing a little arithmetic, we obtain the following summary:

Vocabulary Class		Terms		Percentages	
Name	No.	Number	Frequency	Number	Frequency
Special Terms	3	5	8	1.8	1.5
Allusions	4	0	0	0	0
Key vocabulary	2	142	176	50.7	32.0
One-syllable vocabulary	1	133	366	47.5	66.5
Total,		280	550	100.0	100.0

The number of special terms is down to 5. The number of allusions is down to zero. The total number of words is 550 as compared with 619. The presentation seems to be more logical and much clearer.

Chapter 8, Exhibit 1

SCREENING OUT POTENTIAL TROUBLEMAKERS --
ASSESSMENT OF EACH TERM IN CLASSES 3 AND 4, AND
CERTAIN TERMS IN CLASS 2

Class 3

Term	Frequency	Meaning and Notes	Assessment
AWOL	1	"absent without leave"	acceptable
behavioral principles	1	"principles of human behavior"	probably understandable
communications systems	1	"systems of communications"	acceptable
computers	2	"machines for computing and reasoning"; comments made but not correctly; see Note 1	acceptable
courts martial	1	"courts martial" (there is no adequate synonym)	surely understandable to this audience
cybernetics	1	definition given, though not correctly; see Note 2	acceptable
go over the hill	1	"become absent without leave"	acceptable
naval retraining commands	2	commands for retraining naval personnel	surely understandable to this audience
on board		"included in the Navy"	acceptable
we, our, us	20	many different meanings; see Note 3	ambiguous
personal moral responsibility	1	not clear in the context; see Note 4	probably has different meanings to different people; ambiguous
troublemakers, problem maker	5	men in the Navy who cause disciplinary problems	acceptable, but only the one term "troublemaker" should be used, for the sake of clarity

Note 1, on Computers. The term "computers" occurs in two contexts:

-- "complex mechanical-electronic communications systems, such as computers"

-- "to believe that there is an essential difference between men and highly refined 'computers'"

What is a computer? One authoritative definition is given in the standard international "IFIP-ICC Vocabulary of Information Processing", North Holland Publishing Co., 1966. According to that definition, a computer is:

> A device able to execute a systematic sequence of operations on data, and to perform a substantial computation including numerous arithmetical operations, without the intervention of a human operator.

This definition clearly depends on the definition of "data". The definition of "data" given in the same source is:

> A representation of facts or ideas in a formalized manner, capable of being communicated or manipulated by some process. (Examples: printed text, punched cards, electrical signals.)

The trouble with the use of the term "computer" in the original explanation is that in both occurrences a rather obscure meaning is necessary for it. First, it is not true that every computer is a complex mechanical-electronic communications system (some computers are not electronic at all). Second, there is no indication given as to what meaning "highly refined" is to have in connection with the word "computer".

The term "computer" occurs so often nowadays that many people have rather a good idea of it. It should be fairly clear to this audience.

Note 2 on Cybernetics. This term is defined where it occurs in the explanation. The context is:

> the science of cybernetics (a comparative study of complex mechanical-electronic communications systems, such as computers, with the control system formed by the nervous system and the brain).

Is this definition correct?

Here are some other definitions of "cybernetics":

> -- 1. The comparative study of the control and the internal communication of information-handling machines and the central nervous systems of animals and men, in order to understand better the functioning of

brains and communications. 2. The study of the art of the pilot or steersman.

-- "Glossary of Terms in Computers and Data Processing" by E. C. Berkeley and L. L. Lovett, Berkeley Enterprises, 1960, 96 pp.

-- comparative study of the control system formed by the nervous system and brain and mechanical-electrical communications systems, such as computing machines.

-- Webster's New Collegiate Dictionary, Merriam Webster Co., Springfield, Mass., 1960, 1174 pp.

-- the science of exploring analogies between organic and machine processes

-- IFIP-ICC Vocabulary of Information Processing, op. cit.

-- the science of automatic control

-- Aerospace Dictionary, by Frank Gaynor, Philosophical Library, New York, 1960

-- the comparative study of the automatic control system formed by the nervous system and brain and by mechanoelectrical communication systems and devices (as computing machines, thermostats, photoelectric sorters)

-- Webster's Third New International Dictionary, G. & C. Merriam Co., Springfield, Mass., 1961, 2662 pp.

The modern use of the term originated from a book published in 1948 by John Wiley and Sons, New York, written by the late Norbert Wiener of Mass. Inst. of Technology, which was entitled "Cybernetics, or Control and Communication in the Animal and the Machine".

The term is derived from a Greek word "kybernetes" meaning "steersman, governor."

The term has been used in the years 1948 to the present in many different ways, and has currently rather a diffuse meaning. It is currently an unpopular term with many scientists because of its vagueness.

In regard to the definition of "cybernetics" given in "The Case of the Suspended Sentence", we can comment as follows:

-- "complex" is wrong; simple systems are also included;

-- "mechanical-electronic" is wrong; "mechanical-electrical" would be correct

-- the idea of "the internal automatic control system in the animal and the machine" is not clearly presented in the definition;

-- examples of computers other than thermostats and photoelectric sorters would have helped and should have been given;

-- it would have been desirable to refer to the nervous systems and brains of animals as well as those of human beings.

Note 3, on We, Ours, Us. These terms have been put into class 3 because they shift in meaning. They may actually have any one of eight meanings:

scientists	the Navy
Navy officers	the author of the writing
the United States	the military services of the United States
the author and the audience together	the audience

Following is the evidence:

No.	Context	Meaning
1	"if we could say that human beings ... would automatically react"	scientists
2	"we are not implying that people are in no way predictable"	the author of this writing
3	"nor do we impugn the ... sciences"	the author of this writing
4	"Suppose you said 'We know ....'"	a member of the audience
5	"so let's exclude any recruit"	the Navy

6	"might exclude some of our most successful personnel"	the Navy
7	"is there any simple ... test or question that will tell us..."	the author and the reader together
8	"We know of none to date" (Note: The sense here must be that scientists know of none to date, because clearly the knowledge of the author and the reader together will not be expert enough.)	scientists
9	How big is our manpower supply?	(1) the military services of the United States (2) the Navy
10	Could we leave out all the potential troublemakers...	the Navy
11	we have to have a certain minimum number of these men on board	the Navy
12	As a further example, we find that the man most likely to wind up in a retraining command	(1) the Navy (2) the author and the reader together
13 to 15	We have some of the finest people in the world -- and we have our problem makers	the Navy
16	Justifiable as our complaint may be	the audience
17 to 19	Once we have them, we have to do what we can for their good and the good of the Navy	Navy officers
20, 21	We increase the chances of such success if we catch the potential troublemakers at the beginning	the Navy (because Navy officers by themselves could not do it all)

Note 4, on Personal Moral Responsibility. The term "personal moral responsibility occurs in the statement:

> no machine .... can have a built-in personal moral responsibility

In the Collegiate Dictionary, there are nine definitions of "moral"; no. 7 seems to be the closest:

> sanctioned by, or operative upon, one's conscience or ethical judgment; as, a moral obligation

There are seven definitions of "personal", and No. 1 appears to be the closest:

> of or pertaining to a particular person

Among several definitions of "responsibility" and "responsible" the one that seems to fit closest with the meaning of "responsibility" here is:

> state or quality of being able to respond or answer for one's conduct and obligations

Combining this information, we have the following probable definition for the term:

> quality of being able to answer for one's conduct and fulfil one's obligations as sanctioned by one's conscience or ethical judgment

Fortunately, there are many instances in the experiences of persons where the "personal moral responsibility" of the person is exercised, as for example, a responsibility for treating one's subordinates fairly, with respect, and without favoritism.

So, although this term is vague and hard to define, and although there would be disagreement about which sorts of cases are included and which are not, there is likely to be a common area of meaning. We can accept this term as probably understandable to this audience.

What about the term "built-in"? There is excellent evidence that a personal moral responsibility is not "built-in" for any person. It comes from the training, education, and experience of the young person in the environment (particularly the family) in which he is brought up from a baby to an adult. A person brought up in an environment with no contact with other human beings nor with a human culture will not possess any personal moral responsibility. A pair of eyes may be built into a human being, but not a personal moral responsibility. The term "built-in" is not correct.

Class 4

Term	Freq.	Context or Meaning	Assessment
behavioristic	1	a school of psychological thought	Probably understandable

mechanistic	1	a school of psychological thought	Probably understandable
chi-squares	1	"Chi-squares and coefficients of correlation can ascertain the probabilities of how many people will prefer .... to ...."	Unknown to this audience
coefficients of correlation	1	"	Unknown to this audience
electroencephalograms	1	"Electroencephalograms, psychogalvanometers, and Rorschach diagnoses can tell a lot about what goes on in his mind and in his emotional makeup."	Unknown to this audience
psychogalvanometers	1	"	Unknown to this audience
Rorschach	1	"	Unknown to this audience
fluoroscopes	1	"X-rays and fluoroscopes can strip him of every physiological secret."	Probably understandable
X-rays	1	"	Probably understandable

Class 2

Term	Freq.	Meaning and Notes	Assessments
damnation	1	not clear scientifically; may have objectionable overtones for some readers	to be eliminated
exhausting	1	occurs in the phrase "in exhausting detail"; here it is a wrong word; the phrase should be "in exhaustive detail"	to be changed
full-blooded	1	occurs in the phrase "a full-blooded American boy"; members of a specific race such as Indians can be full-blooded, but very few Americans are; here it is a wrong word; the phrase probably intended is "red-blooded American boy"	to be changed
primrose path	1	taken from a familiar quotation from Shakespeare, "the primrose path of dalliance treads"	probably understandable to this audience

Chapter 9
Practical Aspects of Revising and Improving Explanations

Outline

1. The Value of a Good Explanation
2. The Measurement of a Good Explanation
3. The Effort to Make a Good Explanation
4. The Value of the Computer Program for Vocabulary Analysis
5. Knowledge Necessary for Revising

In this chapter, we take up questions of the practical aspects of good explanation:

- How wortnwhile is a good explanation?
- How much effort is worth putting in, in order to produce good explanation?
- Does it make sense to put a draft explanation into a computer program and analyze its vocabulary? Could you not accomplish more without bothering with the computer?
- How much knowledge does one need in order to revise and improve an explanation?

1. The Value of a Good Explanation

The value of a good explanation is like the value of a key. You may not need a key to go through a door, but if you need a key, and you have the right key, then you can go through and otherwise you can't.

For example, I remember a time when gasoline trucks were labeled "inflammable", though nowadays they are always labeled "flammable". The reason for the change, of

course, is that sometimes the prefix "in-" means "not", as in "ineligible" meaning "not eligible" and "incomplete" meaning "not complete". And some people were confused -- sometimes disastrously confused and lost their lives -- thinking that "inflammable" meant "not able to catch flame". And so they were completely misled by the word "inflammable", believing it to mean the opposite of what it does mean. So the explanation was changed, and the invented word "flammable" is now painted on gasoline trucks.

In such a case as this, the value of the good explanation is enormously high, to the man who might otherwise lose his life. And in a number of cases, if a good explanation makes something possible where a poor one makes it impossible, the good explanation has a similar crucial value.

The value of a good explanation also depends on the number of people who will be using the explanation, and the degree of difference between the good explanation and an ordinary explanation.

Suppose: that 10,000 people will use an explanation; that a good explanation will require an hour to understand; and that an ordinary explanation will require two hours to produce the same degree of understanding. Then, if the time of these people is worth $3 an hour on the average, the difference in value between the two explanations is $30,000. Suppose it takes an extra eight hour day from a man receiving $10 an hour to make the good explanation instead of the ordinary one. Then the investment of $80 in the explainer's time gives a return of $30,000; this is a good bargain.

Of course, these figures are sample figures only. But they show the pattern in which similar calculations could be made for other cases with different data.

2. The Measurement of a Good Explanation

An explanation is good if it produces understanding, if the persons receiving the explanation can do certain desirable things that they could not do previously. For example, they may be able to answer certain questions that they could not answer previously.

Some improvements in explanation require no testing or measuring at all. For example, if in Explanation A, a certain strange term has no definition, while in Explanation B, the strange term does have a good definition, it is undeniable that Explanation B is better, at least in this respect.

Other improvements in explanation may however require testing. It would be normal and reasonable to select a group of persons that were typical of the audience for which the explanation was being prepared. The group could be divided into two equal parts at random; Group 1 and Group 2. Group 1 would read and study Explanation A and answer certain questions; Group 2 would read and study Explanation B and answer the same questions; and comparison of the answers should show differences in the effectiveness of the explanation.

More knowledge of the effectiveness of the explanations could be gained if, after Group 1 had read Explanation A, and answered questions, then they should read Explanation B, and answer the questions. And Group 2 after reading Explanation B, would then read Explanation A, and answer the questions.

3. The Effort to Make a Good Explanation

The amount of work in judging and improving an explanation (as shown in the last two chapters) seems to be enormous. Isn't this amount of work prohibitive?

Actually, it is not. It is the kind of work done regularly by good editors of magazine articles, good writers of technical manuals, and other persons professionally concerned with the presentation of information. The technique once learned takes far less time to carry out than it does to talk about it. It is something like the technique of a good tennis player — much easier to do and do well, than it is to describe it and teach it.

Some of the effort can be delegated to a computer, which can inventory and count the words in the explanation classes, 3, 4, 2, and 1. This focuses the attention of the explainer on what he needs to do.

Also, much of the effort is inevitably made whenever an author or editor changes a first draft of a piece of writing into a second draft. The author or editor considers each term, phrase, and sentence, and tries to clarify and improve the expression of the thought.

4. The Value of the Computer Program for Vocabulary Analysis

How useful is the computer program for vocabulary analysis? Would it not take less effort to apply the principles of revising without using the computer?

Of course, it would be possible to do much the same revising and improving of an explanation without assistance from the computer program.

But the computer program does require or make a decision about each word occurring in the explanation, as to which explanation class it belongs in. Also, when the author examines the lists of words produced by the computer on preliminary runs, he catches a number of mistakes in classifying words and the phrasing of the explanation. The computer focuses a bright searchlight on the vocabulary being used.

The computer program also tells how often words are used. The knowledge that a special term is used many times identifies it as a key idea. The knowledge that a special term is used only once is an invitation to dispense with it.

In one of the original explanations, 29 special terms had a total frequency of 50, an average use of 1.72 times. In its corresponding revised explanation, 22 special terms had a total frequency of 61, an average use of 2.77 times. This change is a clear gain: fewer terms used more often increase understanding and learning by the audience.

The Key Vocabulary consists of the words tagged 2 in the explanation, and they are words of two or more syllables that are used for explaining, that most readers know or probably know, and that are not explained. Some of these words can be replaced with advantage; and possible replacements to be considered could be turned into another guide for explainers.

For example, most of the time (but not always) "often" is just as good as "frequently", "mainly" is better than "primarily", "shorten" is better than "abbreviate", "just right"

is better than "precisely correct", "place" is as good as "situation", etc. We do not wish to change or dilute any desired meaning. But it is silly to use a long word when a short word will do a better job. A guide for translating long words into short words could be helpful.

5. Knowledge Necessary for Revising

What about the amount of knowledge needed to clarify and improve an explanation? Isn't it rather difficult to revise and improve an explanation? Doesn't it require a good deal of technical knowledge?

Certainly, to revise well does require a good deal of technical knowledge, and also a good knowledge of English; and revising is often difficult. In fact revising may be almost impossible if the person revising (the editor) does not have good knowledge of the subject.

However, even in such a case, the editor can:

-- look up terms in a dictionary;

-- read and study the field;

-- consult persons who know a good deal about the field and find out from them what probably is meant by the text.

Sometimes, the editor can take the explanation back to the author and find out from him in many places just what he meant to say. Sometimes, however, the editor completes or perfects what the author was trying to say, what the author meant to say when he actually wrote something different.

Chapter 10

Other Assistance from a Computer in Explaining

Outline

1. Sentence Length
2. Computer Program for Average Number of Words per Sentence
3. Testing of the Program
4. Ratio of Number of Syllables to Number of Words
5. Number of Syllables
6. Classes of Exceptions
7. Omissions
8. Computer Program for Ratio of Number of Syllables to Number of Words
9. Determining a Frequency Distribution and the Sum of Two Frequency Distributions
10. Computer Program
11. Determining the Presence of Key Words and Phrases in an Explanation
12. Computer Program
13. Other Possibilities

In Chapter 6 we described and illustrated the application of a computer program which would take in words and indivisible phrases of a piece of writing, and would classify them in four classes: the Special Vocabulary; the Allusion Vocabulary; the Key Vocabulary; and the One-Syllable Vocabulary. This computer program has been used in a number of places in this report to analyze the vocabulary of various pieces of writing.

Of course there are many other ways also in which a computer program can assist

in the improvement of explanation, and can perform various useful tasks in connection with explanation. Four such ways are described in this chapter, (including stating a program and reporting tests of it.) Many more possible ways can be easily seen.

1. Sentence Length

Among the computer programs that can be useful in controlling the quality of an explanation is one that determines the average number of words per sentence in a piece of writing.

Sentences of an average length of 10 words give a choppy and disorganized impression to the usual adult reader. Sentences with an average length of 18 to 20 words are charactoristic of good standard writing in modern English. Sentences with an average length of over 35 words are rough going. For example, Table 1, which is taken from "The Art of Readable Writing" by Rudolf Flesch (Harper's, 1948) shows the association between the average sentence length (in number of words) and the kind of audience which typically reads writing of that kind. (Of course, since this table is 19 years old, the information in it should be accepted with some caution.)

The average number of words per sentence is, of course, not a sufficient factor to judge the readability of writing. Other important factors affecting readability are: number of syllables per 100 words; number of personal references per 100 words; density of packing of ideas; etc. Even so, it is useful to have a computer program to determine the average sentence length in a piece of writing.

2. Computer Program for Average Number of Words per Sentence

A computer program which determines the average number of words per sentence in a piece of writing is shown in Exhibit 1. This program is written in the computer programming language LISP. This program defines three functions:

NSNS, which is short for "number of sentences"

NWDS, which is short for "number of words"

RAWDSN, which is short for "ratio of number of words to number of sentences"

Table 1

SCALE OF READABILITY OF WRITING

Description of Style of Readability	Average Sentence Length in Words	Average No. of Syllables per 100 Words	Estimated School Grades Completed	Estimated Percent of U.S. Adults
Very Easy	8 or less	123 or less	4th grade	93
Easy	11	131	5th grade	91
Fairly Easy	14	139	6th grade	88
Standard	17	147	7th or 8th grade	83
Fairly Difficult	21	155	some high school	54
Difficult	25	167	high school or some college	33
Very Difficult	29 or more	192 or more	college	4-1/2

Source: "The Art of Readable Writing", by Rudolf Flesch, Harper Brothers, New York, 1949, p. 149

Division in the particular LISP system used discards any remainder. But we would like to have the ratio of number of words to number of syllables to (say) two decimal places; so we define RAWDSN to be 100 times the correct ratio. Then when the answer is reported by the computer, we insert the decimal point two spaces in from the right hand end of the number.

For testing the computer program, we use two examples, PSG1 (short for "Passage 1") which is shown starting at line 00130 in the program, and PSG2 (Passage 2) which is shown starting at line 00200.

3. Testing of the Program

In the lower part of Exhibit 1 appears the machine record of testing of the defined functions on passages 1 and 2.

After the LISP system is in the computer, the user calls the LISP function

"evaluate" using the command EVAL, and gives EVAL the name of the function, "number of sentences of", and the name of the argument, "passage 1", in the style:

EVAL ((NSNS PSG1))

The computer responds 3.

This is correct since passage 1 does have three sentences. (Because of LISP requirements, the sentences are not marked by periods, but are enclosed in parentheses). Similarly, for the other five examples.

The ratio of number of words to number of sentences (the average sentence length) in passage 1 is 15.66 (putting the decimal point back in). This is satisfactorily correct, since the number of words, 47, divided by the number of sentences, 3, is 15.67. It is not necessary in this computer program to be fussy about the failure to round off hundredths of a percent to the nearest hundredth.

4. Ratio of Number of Syllables to Number of Words

Another measure of the difficulty of an explanation is the ratio of the number of syllables to the number of words. If there are many syllables and few words, the explanation is "polysyllabic" and hard to understand. For example, we try to change "Extinguish the illumination" to "Turn out the light".

We desire a computer program which will determine the ratio of the number of syllables to the number of words in a piece of writing.

Such a computer program has been worked out and applies to nearly all cases, producing the correct result in probably more than 99.9% of the cases of words that will actually arise in practice.

This program was also worked out in the computer programming language LISP.

5. Number of Syllables

In general (there are exceptions), we can determine the number of syllables in a word by the following major rule:

Wherever a group of one or more vowels is set off by consonants, a

syllable is produced. Also, if a group of vowels occurs at the end of a word, or at the beginning, one more syllable is produced.

Examples:

R E S P O N S I B I L I T Y : 6 groups of vowels: 6 syllables

B O O K K E E P E R : 3 groups of vowels: 3 syllables

We can express the rule in precise English as a procedure that can be translated into a computer program:

1. Consider each letter of a word successively.
2. Start a counter at 0.
3. If the letter is a vowel and if the next letter is a consonant or there is no next letter, add 1 to the counter.
4. Otherwise, add 0.
5. When the word is finished, the counter will tell the number of syllables.

We can translate this procedure easily into LISP, as follows.

To make sure that the program knows the vowels, we choose VW as a name for vowels and define it in LISP as follows:

```
CSET
(VW (A E I O U Y) )
```

Second, we choose NVCL (an abbreviation for "number of vowel-consonant classes") as a name for the procedure, and express in LISP the following procedure:

1. Define NVCL for any list of letters L as the result of the following program, which makes use of two new variables U and W.
2. Set U equal to the list L at the start.
3. Set W equal to zero at the start.
4. Put down a place marker A.
5. If U is empty, return W as the answer.
6. If the first letter of U is a member of the class of vowels, and if the second letter of U is not a member of the class of vowels, then set W to the old W plus 1. Otherwise, make no change in W.

7. Set U to the old U after dropping the first letter of U.

8. Go to A and repeat.

Exhibit 2 contains some examples of the actual operation of our defined function NVCL.

6. Classes of Exceptions

There are several large classes of exceptions. These are words which end in E, E S, or E D, where the letter preceding this ending is a consonant (examples, C A R E, C A R E S, C A R E D). In such cases we should subtract one from the number given by NVCL.

But there are exceptions to these exceptions. Some are like PUZZLE, PUZZLES, PUZZLED, SABRE, SABRES, SABRED; these have no subtraction. Some are like CHANGES -- no subtraction, but for CHANGE and CHANGED, subtraction. Some are like FADED -- no subtraction, but for FADE and FADES, subtraction.

To take care of them, we shall define a "key consonant class" (abbreviation KCC) which consists of the group of consonants in sequence that precedes the ending E, E S, or E D.

So our large classes of exceptions will be handled by the following rule:

1. Subtract 1 if the word ends in E, E S, or E D, and the preceding letter is a consonant.

2. Add back 1 if the word ends in E and the key consonant class has 2 or more letters and the last letter is L or R (ABLE, SABRE) but the KCC is not L L nor R R (QUADRILLE, FINISTERRE).

3. Add back 1 if the word ends in E S and the KCC (i) has 2 or more letters and the last letter is L or R (CABLES, SABRES) but not L L or R R (real word, QUADRILLES, possible word, FINISTERRES), or (ii) has one or more letters and the last one is C, G, S, X, or Z or the last two letters are C H or S H (BRACES, CAGES, PHRASES, BOXES, BLAZES, PITCHES, WASHES).

4. Add back 1 if the word ends in E D and the KCC has (i) one or more letters and the last one is D or T (FADED, BAITED) or (ii) has 2 or more letters and the last one is L or R (examples, TABLED, SABRED) but not when there are just two letters and they are LL, RL, SL, WL, or RR (CALLED, HURLED, AISLED, BOWLED, STIRRED).

This rule does not provide for short words like HE, YES, and RED, and so we have to insert a condition that the number of syllables is never smaller than 1.

This rule has been expressed in the LISP computer program in Exhibit 2. We shall not try to explain the details here. Examples of its operation are shown.

7. Omissions

The program does not provide for:

-- words like PIQUE and ANALOGUE where the ending is UE yet there is no additional syllable as there is in RETINUE and CONTINUE;

-- words like CREATED, COERCE, COOPERATING, RELIABLE, where two vowels written next to each other produce an extra syllable.

For handling cases like these, which are relatively rare, we could make use of a tag which the computer program would recognize, and then correct the number of syllables. Or we could provide a table of exceptional words which the computer program could refer to. But since the number of syllables is being used as a statistical measure for judging writing, it is more practical to neglect these cases.

8. Computer Program for Ratio of Number of Syllables to Number of Words

Now that we have a way of counting the number of syllables, the rest of the computer program for determining the desired ratio is simple. It closely resembles the preceding program for determining the ratio of the number of words to the number of sentences. (Mathematicians would call the resemblance "isomorphic"

meaning "having precisely the same structure").

This computer program is stated in Exhibit 3. It is tested on each of two passages. One, called Passage 3, is:

> The LISP Language is designed primarily for symbolic data processing.

This has a ratio of 20 syllables to 10 words, or 2.00. The other, called Passage 4, is:

> It has been used for symbolic calculations in differential and integral calculus, electrical circuit theory, mathematical logic, and other fields of artificial intelligence.

This has a ratio of 52 syllables to 23 words, or 2.26. The computer program for reasons described previously gives 100 times this ratio, 200 and 226. On the scale of difficulty of reading given in Rudolf Flesch's book, "The Art of Readable Writing", both these passages score beyond 192 syllables per 100 words, and are therefore classified as "very difficult".

9. Determining a Frequency Distribution and the Sum of Two Frequency Distributions

Another of the programs useful for studying the vocabulary of an explanation is a program which will take in the explanation, and put out a frequency distribution showing each word occurring, and the total number of occurrences of each word in the explanation. The program should also arrange the frequency distribution in alphabetical sequence.

This task has much the same structure as a simpler program which will determine the frequency of occurrence of letters in a list of letters. Of course, there are only 26 different letters whereas there may be hundreds of different words explanation. However, the structure of a program to determine a frequency distribution of letters will be very similar to the structure of a program for determining a frequency distribution of words. Since we are considering structure, we shall consider the simpler problem for letters.

Consider an example which will show the operations that must be expressed in the computer program. Suppose we wish to make a frequency distribution of the string of letters:

T H E L I S P L A N G U A G E

Paying no attention to the order of letters in the alphabet, the first frequency distribution which we obtain is:

T 1	L 2	A 2
H 1	I 1	N 1
E 2	S 1	G 2
	P 1	U 1

Written in a list it is:

T 1 H 1 E 2 L 2 I 1 S 1 P 1 A 1 N 1 G 2 U 1

Now we desire to put this frequency distribution in order according to the alphabetic sequence of letters. The result is:

A 2 E 2 G 2 H 1 I 1 L 2 N 1 P 1 S 1 T 1 U 1

We may have a second string of letters:

I S D E S I G N E D F O R

The alphabetically ordered frequency distribution for this string is:

D 2 E 2 F 1 G 1 I 2 N 1 O 1 R 1 S 2

Combining it with the first frequency distribution, the sum of the frequency distributions is:

A 2 D 2 E 4 F 1 G 3 H 1 I 3 L 2 N 2 O 1 P 1 R 1 S 3 T 1 U 1

These examples display the stages of the process which we desire to express in a computer program.

10. Computer Program

The functions which we need to define in the computer program are the following:

COUNT Given a list of letters, and a particular letter, count the number of its occurrences in the list.

DELETA Given a list of letters, and a particular letter, delete from this list all occurrences of that letter.

FQDS Given a list of letters, produce the frequency distribution consisting of each letter occurring once, and after each letter a number saying how often it occurs.

NUM Given any letter, find its number from 1 to 26 in the alphabet.

SEQ Given a frequency distribution of letters in any order, produce the same frequency distribution but having the letters in alphabetic order.

SUMFD Given two frequency distributions in order produce their appropriate sum in order.

In Exhibit 4 a computer program is stated which defines all these functions, and at the end gives several examples. The last example is:

EVAL ((SUMFD (SEQ (FQDS PSG6)) (SEQ (FQDS PSG5))))

This says:

EVALUATE the sum of the frequency distributions -- namely SUMFD -- found from Passage 6 (PSG6) and from Passage 5 (PSG5).

The answer reported is:

(A 2 D 2 E 4 F 1 G 3 H 1 I 3 L 2 N 2 O 1 P 1 R 1 S 3 T 1 U 1),

which is the same as the sample answer given above.

11. Determining the Presence of Key Words and Phrases in an Explanation

Another thing which we might like to know about an explanation is whether or not certain key words and phrases occur in it.

As an illustration, suppose that we wanted to know whether any of a group of words or phrases that mark the presence of an example or instance occur in an explanation. The list might be:

such as	say	for example
for instance	for an example	

12. Computer Program

A computer program has been worked out in LISP which will report whether or not any one of a list of given words and phrases occurs in an explanation.

This computer program is stated in Exhibit 5.

In order to use this program, the following procedure applies:

1. Make a list of the key words and phrases that we desire to check; call it L.
2. Give this list to LISP in the form of a list of sublists of atoms (LISP term), each atom being an English word.
3. Give the passage which we wish to examine to LISP in the form of a list of atoms (LISP term), each of which is an English word. Call it P.
4. Employ the defined LISP function OCUR, which is short for "report whether any one of the key words or phrase OCcURs in the passage", with two arguments L and P.
5. If one occurs, the report is *T* (LISP term for "true").
6. If none occurs, the report is NIL (LISP term for "false").

This program applies to phrases of single words, phrases of two consecutive words, and phrases of three consecutive words.

This program can be easily modified if phrases of four or more words are to be checked. It can also be easily modified to cover cases where a phrase may be divided. For example, consider "such as" which may be separated as in the passage: "Corundum, which is aluminum oxide, occurs in nature in such gem varieties as ruby and sapphire, and as the common, economically important mineral emery, useful for abrasives."

Illustrations of the use of this program are shown in Appendix 5. Appendix 5 shows that a list of five key words and phrases is given to LISP, with the command CSET (which may be thought of as meaning "CONSTANT SET"); this list is named LS1,

which is short for "List 1". Then five short passages are also given to LISP with the command CSET; their names are PSG8, PSG9, PSG10, PSG11, and PSG12. One of them, passage 10, contains none of the key words or phrases. The other four all contain different key words or phrases.

The testing of the program on these examples is also shown. The LISP function EVAL is called, and EVAL is given the testing function OCUR and its two arguments, L and P. L in this case is only LS1. P has five possible values PSG8, PSG9, PSG10, PSG11, and PSG12. The answers *T* (for "true") and NIL (for false) are correct, and the program therefore appears to be correct.

Exhibit 1

COMPUTER PROGRAM FOR AVERAGE NUMBER OF WORDS PER SENTE"CE

```
00010    DEFINE ((
00020    (NSNS (LAMBDA (L) (LENGTH L)  ))
00030    (NWDS (LAMBDA (L) (COND
00040       ((NULL L) 0)
00050       (T (PLUS (LENGTH (CAR L)) (NWDS (CDR L))))    )))
00060    (RANDSN (LAMBDA (L) (QUOTIENT (TIMES 100 (NWDS L))
00070              (NSNS L))  ))
00110    ))
00120    CSET
00130    (PSG1 ((THE PURPOSE OF THIS  BOOK IS TO PROVIDE AN INTRODUCTION
00140      TO TEACHING MACHINES) (THIS BRANCH OF KNOWLEDGE HAS GROWN
00150      OUT OF PHYSICAL EXPERIMENTS DEALING WITH THE PROCESSES
00160       OF LEARNING AND TEACHING) (THESE EXPERIMENTS HAVE INCLUDED
00170       VARIED SUBJECTS SUCH AS PIGEONS AND RATS AS WELL AS HUMAN
00180      BEINGS)) )
00190    CSET
00200    (PSG2((THE LISP LANGUAGE IS DESIGNED PRIMARILY FOR SYMBOLIC
00210       DATA PROCESSING) (IT HAS BEEN USED FOR SYMBOLIC CALCULATIONS
00220        IN DIFFERENTIAL AND INTEGRAL CALCULUS ELECTRICAL CIRCUIT THEORY
00230        MATHEMATICAL LOGIC AND OTHER FIELDS OF ARTIFICIAL INTELLIGENCE)
00240        (LISP IS A FORMAL MATHEMATICAL LANGUAGE))  )
00250    STOP)))) )))))
```

```
✓EVAL ((NSNS PSG1))
   3
✓EVAL ((NWDS PSG1))
   47
✓ EVAL ((NWDS PSG2))
   39
✓ EVAL ((NSNS PSG2))
   3
✓ EVAL ((RANDSN PSG1))
   1566
✓ EVAL ((RANDSN PSG2))
   1300
```

Exhibit 2

COMPUTER PROGRAM FOR NUMBER OF SYLLABLES

```
00010    DEFINE ((
00020    (NVCL (LAMBDA (L) (PRØG (U W)
00030        (SETQ U L) (SETQ W 0)
00040        A (CØND ((NULL U) (RETURN W))
00050            ((AND (MEMBER (CAR U) VW)
00060                  (ØR (NULL (CDR U))
00070                      (NØT (MEMBER (CADR U) VW)))) (SETQ W (ADD1 W))))
00080            (SETQ U (CDR U))  (GØ A)  )))
00090    ))
00110    DEFINE ((
00120    (SYCR (LAMBDA (L) (CØND
00130       ((AND (GREATERP (NVCL L) 1) (APSØ L))
00140          (PLUS  (MINUS 1) (CØRR2 L) (CØRR3 L) (CØRR4 L)))
00150        (T 0)   )))
00160    (APSØ (LAMBDA (L) (PRØG (U)
00170        (SETQ U (REVERSE L))
00180        B (CØND ((NULL U) (RETURN NIL))
00190                ((NULL (CDR U)) (RETURN NIL))
00200                ((NULL (CDDR U)) (RETURN NIL)))
00210        A (CØND ((AND (EQ (CAR U) (QUØTE E))
00220         (NØT (EQØN (CADR U) VW))) (GØ C))
00230         ((AND (ØR (EQ (CAR U) (QUØTE D))
00240          (EQ (CAR U) (QUØTE S)))
00250         (EQ (CADR U) (QUØTE E))
00260         (NØT (EQØN (CADDR U) VW))) (GØ C)))
00270         (RETURN NIL)
00280       C (RETURN T) )))
00290    (NSYL (LAMBDA (L) (PLUS (NVCL L) (SYCR L))))
00300    ))
00320    DEFINE ((
00330    (EQØN (LAMBDA (L M) (CØND
00340      ((NULL M) NIL)  ((EQUAL L (CAR M)) T)
00350      (T (EQØN L (CDR M))) )))
00360    (LAST (LAMBDA (L) (CØND
00370      ((NULL L) NIL)
00380      ((NULL (CDR L)) (CAR L))
00390      (T (LAST (CDR L))) )))
00400    (H2ME (LAMBDA (L) (CØND
00410      ((NULL L) NIL)  ((NULL (CDR L)) NIL)  (T T)  ) ))
00420    ))
00430    CSET (VW (A E I Ø U Y))
00440    DEFINE ((
00450    (CØRR1 (LAMBDA (L) (PRØG (U)
00460       (SETQ U (REVERSE L))
00470       A (CØND ((AND (EQ (CAR U) (QUØTE E))
00480          (NØT (EQØN (CADR U) VW))) (GØ C))
00490          ((AND (ØR (EQ (CAR U) (QUØTE D))
00500           (EQ (CAR U) (QUØTE S)))
00510         (EQ (CADR U) (QUØTE E))
00520        (NØT (EQØN (CADDR U) VW))) (GØ C)))
00530        (RETURN 0)
00540       C (RETURN 1) )))
00550    (CØRR2 (LAMBDA (L) (PRØG (U X)
00560       (SETQ U (REVERSE L)) (SETQ X (KCC L))
00570    (CØND ((AND (EQ (CAR U) (QUØTE E))
00580           (H2ME X) (EQØN (LAST X) (QUØTE (L R)))
00590           (NØT (EQØN (LSTW X) (QUØTE ((L L ) (R R))))))
00600           (RETURN 1)))
00610        (RETURN 0)  )))
```

```
00620    (NXLS (LAMBDA (L) (COND
00630       ((NULL L) NIL) ((NULL (CDR L) ) NIL)
00640       ((NULL (CDDR L)) (CAR L))
00650       (T (NXLS (CDR L)))   )))
00660    (CORR3 (LAMBDA (L) (PROG (U W X)
00670       (SETQ X (KCC L))
00680      (SETQ U (REVERSE L)) (SETQ W 0)
00690       (COND ((AND (EQ (CAR U) (QUOTE S))
00700         (EQ (CADR U) (QUOTE E))
00710         (OR (AND (H2ME X) (EQON (LAST X) (QUOTE (L R))))
00720          (EQON (LAST X) (QUOTE (C G S X Z)))
00730          (AND (EQ (LAST X) (QUOTE H))
00740          (EQON (NXLS X) (QUOTE (C S))))))
00750       (SETQ W  1)))
00760       (RETURN W) )))
00770    (CORR4 (LAMBDA (L) (PROG (U X)
00780      (SETQ U  (REVERSE L)) (SETQ X (KCC L))
00790        (COND ((AND (EQ (CAR U) (QUOTE D))
00800        (EQ (CADR U) (QUOTE E))
00810        (OR (EQON (LAST X) (QUOTE (D T)))
00820        (AND (EQON (LAST X) (QUOTE (L R)))
00830        (NOT (EQON (LSTW X) (QUOTE ((L L) (R L) (S L)
00840           (W L) (R R)))) ))))
00850       (RETURN 1)))
00860        (RETURN 0)  )))
00870    (LSTW (LAMBDA (L) (LIST (NXLS L) (LAST L))))
00880    (APS (LAMBDA (L) (PROG (U)
00890       (SETQ U (REVERSE L))
00900       (COND ((OR (AND (EQ (CAR U) (QUOTE E))
00910                       (NOT (EQON (CADR U) VW)))
00920                  (AND (EQ (CADR U) (QUOTE E))
00930                       (OR (EQ (CAR U) (QUOTE S))
00940                           (EQ (CAR U) (QUOTE D)))
00950                       (NOT (EQON (CADDR L) VW))))     (RETURN T)))
00960       (RETURN NIL)   )))
00970    (KCC (LAMBDA (L) (PROG (U W)
00980       (SETQ U (REVERSE L)) (SETQ W NIL)
00990    C1  (COND ((NULL U) (GO A5))
01000     ((EQ (CAR U) (QUOTE E)) (GO A1)))
01010       (GO C2)
01020    C2  (COND ((NULL (CDR U)) (GO A5))
01030             ((AND (EQ (CADR U) (QUOTE E))
01040            (OR (EQ (CAR U) (QUOTE S))
01050                (EQ (CAR U) (QUOTE D)))) (GO A2)))
01060       (GO A5)
01070    A1 (SETQ U (CDR U)) (GO A3)
01080    A2 (SETQ U (CDDR U)) (GO A3)
01090    A3 (COND ((NULL U) (GO A5))
01100           ((EQON (CAR U) VW) (GO A5)))
01110         (SETQ W (CONS (CAR U) W))
01120          (SETQ U (CDR U)) (GO A3)
01130    A5 (RETURN W)    )))
01140    ))
01150    STOP))))) ))))))
```

```
✓NVCL ((R E S P Ø N D))
  2
✓NVCL ((R E S P Ø N S I B I L I T Y))
  6
✓NVCL ((A R E))
  2
✓NVCL ((A D M I R A B L E))
  4


 NSYL ((H U R L E D))
  1
 NSYL ((C H U R C H E S))
  2
 NSYL ((F A D E S))
  1
 NSYL ((R E S P Ø N S I B I L I T Y))
  6
 NSYL ((A C C I D E N T))
  3
 NSYL ((B E R T H))
  1
 NSYL ((P E C U N I A R I L Y))
  5
✓NSYL ((P U Z Z L E D))
  2
✓NSYL ((H U L L E D))
  1
 NSYL ((T A B L E D))
  2
 NSYL ((B L E D))
  1
 NSYL ((E D))
  1
```

Exhibit 3

COMPUTER PROGRAM FOR RATIO OF NUMBER OF SYLLABLES TO NUMBER OF WORDS

```
DEFINE ((
(NSYLS (LAMBDA (L) (COND
   ((NULL L) 0)
   (T (PLUS (NSYL (CAR L)) (NSYLS (CDR L))))  )))
(NWDS (LAMBDA (L) (LENGTH L)  ))
(RASYWD (LAMBDA (L) (QUOTIENT (TIMES 100 (NSYLS L))
      (NWDS L))  ))
))
CSET
(PSG3 ((T H E) (L I S P) (L A N G U A G E) (I S) (D E S
    I G N E D) (P R I M A R I L Y) (F O R) (S Y M B O L I C)
    (D A T A) (P R O C E S S I N G)) )
CSET
(PSG4 ((I T) (H A S) (B E E N) (U S E D) (F O R) (S Y M B O L I C)
    (C A L C U L A T I O N S) ( I N) (D I F F E R E N T I A L)
    (A N D) (I N T E G R A L) (C A L C U L U S) (E L E C T R I
    C A L) (C I R C U I T) (T H E O R Y) (M A T H E M A T I C A L)
    (L O G I C) (A N D) (O T H E R) (F I E L D S) (O F)
    (A R T I F I C I A L) (I N T E L L I G E N C E)) )
STOP)))))  )))))
```

```
✓EVAL ((NWDS PSG3))
  10
✓ EVAL ((NSYLS PSG3))
  20
✓ EVAL ((NWDS PSG4))
  23
✓ EVAL ((NSYLS PSG4))
  52
✓ EVAL ((RASYWD PSG3))
  200
✓ EVAL ((RASYWD PSG4))
  226
```

Exhibit 4

COMPUTER PROGRAM FOR FREQUENCY DISTRIBUTION OPERATIONS

```
DEFINE ((
(DELETA (LAMBDA (X L) (PROG ( U V)
  (SETQ U L) (SETQ V NIL)
  A ( COND ((NULL U) (RETURN V))
  ((NULL (EQUAL X (CAR U))) (SETQ V (APPEND V (LIST (CAR U))))))
  (SETQ U (CDR U)) (GO A)) ))
(FQDS (LAMBDA (L) (PROG (U V)
  (SETQ U L) (SETQ V NIL)
  A (COND ((NULL U) (RETURN  V)))
  (SETQ V (APPEND V (LIST (CAR U) (COUNT (CAR U) U))))
  (SETQ U (DELETA (CAR U) U))
  (GO A) )))

(COUNT (LAMBDA (X L) (PROG (U V)
  (SETQ U L) (SETQ V 0)
  A (COND ((NULL U) (RETURN V))
    ((EQUAL (CAR U) X) (SETQ V (PLUS 1 V))))
  (SETQ U (CDR U))  (GO A)  )))
))
DEFLIST ((
(A 1) (B 2) (C 3) (D 4) (E 5) (F 6) (G 7) (H 8) (I 9) (J 10) (K 11)
(L 12) (M 13) (N 14) (O 15) (P 16) (Q 17) (R 18) (S 19) (T 20) (U 21)
(V 22) (W 23) (X 24) (Y 25) (Z 26)
) RANK )
DEFINE ((
(NUM (LAMBDA (X) (GET X (QUOTE RANK))))

(SEQ (LAMBDA (L) (PROG (U V #)
   (SETQ U L) (SETQ V NIL) (SETQ # 1)
    A (COND ((NULL U) (GO B))
      ((EQUAL (NUM (CAR U)) #) (SETQ V
          (APPEND V (LIST (CAR U) (CADR U))) )))
    (SETQ U (CDDR U)) (GO A)
    B (COND ((EQUAL # 26) (RETURN V)))
     (SETQ # (ADD1 #))
       (SETQ U L) (GO A)   )))
))
CSET
(PSGS (T H E L I S P L A N G U A G E))
```

```
DEFINE ((
(SUMFD (LAMBDA (L M) (PRØG (U V W)
   (SETQ U L) (SETQ V M) (SETQ W NIL)
C (CØND ((NULL U) (GØ A2))
      ((NULL V) (GØ A3))
      ((EQUAL (CAR U) (CAR V)) (GØ A6))
      ((SMLP (CAR U) (CAR V)) (GØ A4))
      ((SMLP (CAR V) (CAR U)) (GØ A5))
      (T (RETURN (QUØTE HELP)))  )
A2 (SETQ W (APPEND W V)) (RETURN W)
A3 (SETQ W (APPEND W U)) (RETURN W)
A6 (SETQ W (APPEND W (LIST (CAR U) (PLUS (CADR U)
       (CADR V) )) ))
      (SETQ U (CDDR U)) (SETQ V (CDDR V)) (GØ C)
A4 (SETQ W (APPEND W (LIST (CAR U) (CADR U))) )
      (SETQ U (CDDR U)) (GØ C)
A5 (SETQ W (APPEND W (LIST (CAR V) (CADR V))) )
      (SETQ V (CDDR V)) (GØ C)    )))
(SMLP (LAMBDA (X Y) (CØND
   ((LESSP (NUM X) (NUM Y)) T) (T NIL) )))
))


(SET (PSG6 (I S D E S I G N E D F Ø R))
(SET (PSG7 (S Y M B Ø L I C D A T A P R Ø C E S S I N G))


STØP)))))  )))))


✓EVAL ((SEQ (FQDS PSG7)))
   (A 2 B 1 C 2 D 1 E 1 G 1 I 2 L 1 M 1 N 1 Ø 2 P 1 R 1 S 3 T
   1 Y 1)
✓EVAL ((SUMFD (SEQ (FQDS PSG6)) (SEQ (FQDS PSG5))))
   (A 2 D 2 E 4 F 1 G 3 H 1 I 3 L 2 N 2 Ø 1 P 1 R 1 S 3 T 1 U
   1)
```

Exhibit 5

COMPUTER PROGRAM FOR DETERMINING PRESENCE OF SPECIFIED WORDS AND PHRASES

```
DEFINE ((
(OCR (LAMBDA (L P) (COND
    ((AND (MEMBER (CAR L) P)
        (OR (NULL (CDR L))
           (AND ( EQ (CADR L) (GETB (CAR L) P))
            (OR (NULL (CDDR L))
              (EQ (CADDR L) (GETB (CADR L) P))))))  T)
    (T NIL) )))
(GETB (LAMBDA (X L) (COND
    ((NULL L) NIL)
    ((EQ X (CAR L)) (CADR L))
    (T (GETB X (CDR L))) )))
(OCUR (LAMBDA (L P) (PROG (U W)
    (SETQ U L) (SETQ W NIL)
    A (COND ((NULL U) (RETURN W))
        ((EQ W T) (RETURN T))
      ((OCR (CAR U) P) (SETQ W T)) )
      (SETQ U (CDR U)) (GO A) )))
))

 CSET (
 LS1 ((SUCH AS) (FOR INSTANCE) (FOR EXAMPLE) (SAY) (FOR AN EXAMPLE) ) )
 ( .2 (
 PSG8 (THE MATHEMATICAL ABILITIES OF THE COMPUTER ARE VERY
 USEFUL TO DESIGN ENGINEERS  THE WING DESIGN OF A SUPERSONIC
 PLANE FOR EXAMPLE DEPENDS ON MANY FACTORS  THE DESIGNER
 DESCRIBES EACH OF THESE FACTORS MATHEMATICALLY IN THE FORM OF
 AN EQUATION))
 CSET (
 PSG9 (MANY FACTORS SUCH AS SPEED AND WEIGHT AFFECT
 THE DESIGN OF PLANES))
 CSET (
 PSG10 (THE DESIGNER CAN USE PENCIL AND PAPER OR A HALF
 HOUR VISIT TO THE COMPUTER ROOM))
 CSET (
 PSG11 (THE DESIGNER CAN FOR AN EXAMPLE USE A DRAWING
 ON GRAPH PAPER))
 CSET (
 PSG12 (THE DESIGNER FOR  NSTANCE COULD CHOOSE THE TYPE OF
 COMPUTER TO ASSIST HIM))

                    ✓EVAL ((OCUR LS1 PSG9))
                      *T*
                    ✓EVAL ((OCUR LS1 PSG8))
                      *T*

                    ✓EVAL ((OCUR LS1 PSG10))
                      NIL
                    ✓EVAL ((OCUR LS1 PSG12))
                      *T*

                    ✓EVAL ((OCUR LS1 PSG11))
                      *T*
```

Chapter 11

A Short Summary of Advice for Explaining

If you desire to produce good explanation, what are the most important rules that you should keep in mind?

Different people of course would answer this question in different ways; but here is what seems to me to be a reasonable answer.

1. Audience. Consider your audience all the time. These are the people who you hope will understand what you want to tell them. They are looking at their wrist watches and eager to go away. So you must interest them, and persuade them to listen.

2. Subject Matter. Have something worth saying, and say it plainly. A lot of changing and revising may be needed until finally you attain something worth saying on a particular subject to a particular audience.

3. Understanding. The purpose of an explanation is to produce understanding in an audience. If understanding is produced, the real world changes: the persons in the audience become able to make new choices, to do new things, to act differently. Whenever possible, judge and improve explanations by testing the understanding produced in the audiences.

4. Common Properties and Relations. When explaining a new idea, report its common properties and relations. In this way, you answer the natural questions that an audience will ask about a new term. And in this way your audience can fasten new information on to old knowledge with a clear and reasonable association.

5. Vocabulary. The words you use, their meanings and their overtones, are crucially important. Always be aware of the meanings of words to your audience. And make sure you give a definition (or some defining statements) using words they understand, for every special term or difficult term that you need to introduce.

6. Personal Attitude. The explainer's personal attitude towards his audience has a big influence on the understanding produced. The relation of the explainer to his audience should be very close to friendship. For instance, when Dr. Spock says in his book on child care (see the bibliography): "If a child is so carefully watched that he never has an accident, he is being fussed over too much; his bones may be saved, but his character will be ruined", you can feel the bond of friendship and sympathy between Dr. Spock and his audience.

7. Stopping. It is not necessary to explain every last detail you can think of; many good explanations have been spoiled by saying too much. As soon as you have said what is needed, stop.

Appendix 1

Computer Program for Vocabulary Classification Using a DEC PDP-7 Computer

This appendix specifies a program which has been denoted EXPL 5, standing for "Explanation Program Edition 5". It is expressed in symbolic assembly language for the PDP-7 Computer made by Digital Equipment Corporation, Maynard, Mass. The program will sort the vocabulary of a piece of writing into the four explanation classes (numbered 3, 4, 2, 1, and described above in Chapter 5 of this report).

Edition 5 of the program here given has been superseded slightly, by Edition 6, but Edition 6 at present writing has not been completely debugged, and so is not given here.

Most of this appendix is written only for a limited audience: those persons who want to thoroughly understand the program and who are thoroughly familiar with programming and with the PDP-7 computer. Most of this appendix is not written for the much larger audience assumed for the rest of this report.

The operating instructions for the binary program (produced by assembling the symbolic program here stated) are as follows:

1. Load DDT (DEC Debugging Tape). Give the commands KILL$, ZERO$.
2. Place the EXPL 5 binary program tape in the tape reader.
3. Give DDT the command DEBUG$. (The binary program tape will load placing its symbols into DDT's symbol table).
4. Give DDT the command BEG' (which will start the program EXPL 5 at register BEG = 2744 octal, listening for commands).

5. The programmed computer is now waiting for text to be inserted and summarized. To insert text, type I (for "insert"), then the text; for example; IHERE IS SOME EXPLANATION TO BE SORTED.$$
At the finish of the text, press the key ALT MOD twice (this action prints "$"). The inserted text will be moved from the Command Buffer to the Storage Buffer, and at the end the typewriter will give a carriage return and line feed.

6. To verify that the text is contained in the computer, type P, ALT MOD, ALT MOD; this action will show as:

P$$

Then the text will print:

HERE IS SOME WRITING TO BE SORTED.

For other commands, see Chapter 6 in which this program is further described.

The maximum length of text which will be accepted on any single insert command is 1000 characters, or about 15 lines of 60 characters each. The program will be clobbered if more than 1000 characters are inserted at any one insert command.

The maximum total length of text which will be stored at one time for analysis inside the computer is 4000 characters.

Note: Since the CAL (subroutine call) instruction in the PDP-7 computer has the machine language code 00, it is regularly omitted in the computer program. Hence, the name of a subroutine by itself may be used as a macro calling that subroutine, and when finished it automatically returns control to the routine which called it. Macros may also be constructed in other ways.

Outline

Section	Title
1	The Computer Program for Vocabulary Classification, expressed in Symbolic Assembly Language
2	The Subroutine Symbols in Alphabetic Order
3	The Subroutine Symbols in Numeric Order
4	Mnemonic Key to the Subroutine Symbols
5	Typewriter Character Codes (ASCII)

THE COMPUTER PROGRAM

```
/EXPL EDIT 5 8/25/66
CAC=6
$1/       DAC CAC
          LAC 20
          ISZ P
          DAC I P
          AND (17777)
          TAD (JMP I-1)
          DAC .+2
          LAC CAC
          JMP I .

RZ,       CLA!SKP
X1,       ISZ I P
X,        DAC CAC
          LAM
          TAD P
          DAC P
          TAD (JMP I 1)
          DAC .+2
          LAC CAC
          JMP I .

P,        0
PDL,      PDL+100/

PUSH=JMS .
          0
          ISZ P
          DAC I P
          JMP I PUSH-JMS

POP=JMS .
          0
          LAM
          TAD P
          DAC P
          DAC 17
          LAC I 17
          JMP I POP-JMS

CIA,      CMA
          TAD (1)
          JMP X

SEAR,     CIA
          SMA
          JMP X
          DAC CNTR
          LAC CPTR
          DAC CPTRS
SE1,      LAC PT
          DAC SPT
          LAC CPTRS
          DAC CPTR
SE2,      RCH
          SAD (175)
          JMP SFD
          DAC TARG
SE4,      SAD (43)
          JMP MATN
          LAC (SAD TARG)
SE3,      DAC SEM1
          LAC SPT
          SAD Z
          JMP SNFD
          ISZ SPT
          GETN
SEM1,     XX
          JMP SE2
          ISZ PT
          JMP SE1
SFD,      LAC SPT
          DAC PT
          ISZ CNTR
          JMP SE1
          JMP X

SNFD,     DZM PT
          ERR

PINCH,    LAM-144
          FEED
          LAW PPA
          OUTX
          LAM-5
          FEED
          LAW 14
          PPA
          LAM-3
          DAC CNTR
          LAW 377
          FEED1
          LAM-500
FEED,     DAC CNTR
          CLA
FEED1,    PPA1
          ISZ CNTR
          JMP .-2
          JMP X
```

```
PPA,      AND (177)
          XOR (200)
PPA1,     PSA
          PSF
          JMP .-1
          JMP X+1
SPT,      0
READ,     DZM PT
          DZM Z
APPEND,   RSF
          RSA
READ1,    RPA
          SAD (14)
          JMP X
          DAC TARG
          LAC Z
          ISZ Z
          SETIN
          LAC TARG
          PUT
          JMP READ1

RPA,      LAW
          DAC CNTR
RPA2,     RSF
          JMP RPA1
          RRB!RSA
RPA3,     AND (177)
          SZA
          SAD (177)
          JMP RPA
          JMP X
RPA1,     ISZ CNTR
          JMP RPA2
          RRB
          LAW 14
          JMP RPA3

MATN,     DZM MAT1
          LAC (JMP MAT2)
          JMP SE3
MAT1,     0
MAT2,     TLETN
          JMP MAT3
          CLC
          DAC MAT1
          JMP SE4
MAT3,     ISZ MAT1
          JMP SE4
          JMP SC2

PROC,     LAC Z
          SETIN
          CLA
          PUT
          DZM OLDZ
          LAM-1776+1
          DAC CNTR
          LAW HASHT-1
          DAC 10
          DZM I 10
          ISZ CNTR
          JMP .-2

RWD,      LAC OLDZ
          DAC IN
          DZM HASHAD
          DZM CHRCT
RWL,      LAC OLDZ
          SAD Z
          JMP PROUT
          ISZ OLDZ
          GTLETN
          JMP WRDE
          ALSS 5
          ADD HASHAD
          IDIVS I
          777
          DAC HASHAD
          ISZ CHRCT
          JMP RWL

WRDE,     LAC CHRCT
          SNA
          JMP RWD
          SAD (1)
          JMP TTAG
ORD,      LAC OLDZ
          DAC OLDZS
          ES
          JMP ENT
          ISZ I LAST
          JMP ENT1

ENT,      LAC IN
          XOR (400000)
          DAC I LAST
          ISZ LAST
          VOCL
          DAC I LAST
ENT1,     LAC OLDZS
          DAC OLDZ
          JMP RWD
OLDZS,    0
```

```
GTLETN, GETN
TLETN,  SAD (55)
        JMP YESLN
        TAD (-60+1)
        SPA
        JMP X+1
        TAD (-12+1)
        SPA
        JMP YESLN
        TAD (-101+72+1)
        SPA
        JMP X+1
        TAD (-32+1)
        SPA
YESLN,  ISZ I P
        JMP X+1

TTAG,   LAC IN
        GETN
        SAD (101)
        JMP ORD
        SAD (111)
        JMP ORD
        SAD (117)
        JMP ORD
        TAD (-100+1)
        SPA!CLL
        JMP NUMTAG
        TAD (40)
        ALS 6
        DAC TAG
        LAC (7700)
        JMP TAGA

NUMTAG, TAD (40)
        ALSS 14
        DAC TAG
        LAW 10000
TAGA,   AND I LAST
TAGOK,  XOR TAG
        XOR I LAST
        DAC I LAST
        JMP RWD

C2W,    DAC OLDZ
        LAC IN
C2WN,   DAC IN1
C2WL,   LAC OLDZ
        ISZ OLDZ
        GTLETN
        JMP C2W1
        DAC TAG
        LAC IN1
        ISZ IN1
        GTLETN
        JMP .+2
        CMA
        TAD (1)
        TAD TAG
        SZA
        JMP X
        JMP C2WL
C2W1,   LAC IN1
        ISZ IN1
        GTLETN
        JMP RZ
        CLC
        JMP X
C2WLOW, LAC LOW
        JMP C2WN

ES,     LAC HASHAD
        ALSS 1
        ADD (HASHT
ESL,    DAC LAST
        LAC I LAST
        SNA
        JMP X
        C2W
        SNA
        JMP ESF
        LAC LAST
        TAD (2)
        SAD (HASHT+1776)
        LAC (HASHT)
        JMP ESL
ESF,    ISZ LAST
        JMP X1

HASHT,  HASHT+1776/
-0      -0
HASHAD, 0
IN1,    0
TAG,    0
LAST,   0
LOW,    0
TARG,   0
HIGH,   0
CLASS,  0
FTF,    0
```

```
PROUT,  LAW ATAB-1
        DAC 10
        LAC Z
        DAC LOW
PR1,    LAC Z
        SAD LOW
        JMP PR1A
        LAC (566500)
        TYW
        LAC NU
        DPT
        TSP
        LAC (466100)
        TYW
        LAC FQ
        DPT
PR1A,   DZM NU
        DZM FQ
        LAC I 10
        SNA
        JMP CRLF
        DAC CLASS
        LAC Z
PR2,    DAC LOW
        CLC
        DAC FTF
        LAW HASHT-1
        DAC 11
PR3,    LAC I 11
        SAD (-0)
        JMP PR4
        DAC TARG
        DAC OLDZ
        LAC I 11
        LRSS 14
        SAD CLASS
        JMP PR5
        JMP PR3

PR5,    C2WLOW
        SPA!SNA
        JMP PR3
        LAC TARG
        ISZ FTF
        JMP CHIGH
PR6,    LAC TARG
        DAC IN
        LAC 11
        DAC HIGH
        JMP PR3

NU,     0
FQ,     0
```

```
CHIGH,  C2W
        SMA
        JMP PR3
        JMP PR6

PR4,    ISZ FTF
        JMP PR7
        JMP PR1
PR7,    LAC LOW
        SAD Z
        PR8
PR9,    LAC HIGH
        PRDAT
        CRLF
        LAC IN
        JMP PR2
```

```
PR8,    CRLF
        CRLF
        LAC (435441)
        TYW
        LAC (636300)
        TYW
        LAC CLASS
        PRSIX
        JMP CRLF

PRDAT,  ADD (-2)
        DAC 12
        LAC I 12
        DAC TARG
PRD2,   GTLETN
        JMP PRD1
        TYO
        ISZ TARG
        LAC TARG
        JMP PRD2
PRD1,   TSP
        LAM-12
        ADD TBC
        SPA
        JMP PRD1
        LAC I 12
        DAC TARG
        AND (77)
        TAD (1)
        LMQ
        ADD FQ
        DAC FQ
        ISZ NU
        LACQ
        DPT
        TSP
        LAC TARG
        LRS 6
        AND (77)
        SZA
        JMP PRSIX
        JMP X

TSP,    CLA ISKP
T3S,    CLA ISKP
TSPA,   XOR (404000)
TYW,    LMQ
        TYWI
        TYWI
TYW1,   LLS 1006
        TAD (40)
        SAD (100)
        JMP X
        JMP TYO

PRSIX,  AND (77)
        JMP TSPA

ATAB,   23          24          22          21          0
```

```
DEG,    LAC (POL-1)
        DAC P
LIS,    CRLF
        LAC (CBUF)
        DAC IN
        DAC CPTR
LI1,    LAM-1
        DAC ALTF
LI2,    TYI
        SAD (177)
        JMP RUB
        PUTI
        SAD (175)
        JMP ALT
        JMP LI1
RUB,    LAC IN
        SAD (CBUF)
        JMP LIS
        ADD (377777)
        DAC IN
        GETAC
        TYO
        JMP LI1
ALT,    ISZ ALTF
        JMP LI2
        CRLF
        LAC IN
        DAC COMEND
CD,     DZM NUM
        DZM ARG
PLUS,   LAC (ADD SYL)
NEWOP,  DAC OP
        DZM SYL
CDR,    RCH
        TAD (-37)
        SPA
        JMP CD
        TAD (-77)
        SMA
        JMP LIS
        TAD (LAC DTB+100)
        DAC CDG
CDG,    XX
        SMA
        ERR
        AND (1-1)
        DAC CDX
        LAC I CDG
        RAL
        SPA!RAL
        JMP I CDX
        SPA!CLA
        LAC (JMP)
        XOR CDX
        DAC CDX
OP1,    LAC NUM
OP,     XX
        DAC NUM
        LAC ARG
        RAL
        LAC NUM
        SNL
        JMP CDY
OP2,    SAD (-0)
        CLA
CDX,    XX
        JMP CD
RCH,    LAC CPTR
        SAD COMEND
        ERR
        GETAC
        LAC (400000)
        ADD CPTR
        DAC CPTR
        JMP X+1

CDY,    LAC OP
        RTL
        LAM-1
        SZL!CLL
        CMA
        JMP OP2
CPTR,   0
ALTF,   0
COMEND, 0
ARG,    0
NUM,    0
SYL,    0
DIG,    LAC SYL
        MUL S I
        12
        LACQ
        TAD CDG
        TAD (-LAC-DTB-20+1)
        DAC SYL
        CLC
        DAC ARG
        JMP CDR
```

```
CBUF,   CBUF+1000,
PUT,    LAC IN
        SPA
        JMP PUT1
        LAC I IN
        LRS 11
        LAC CAC
        LLS 11
        DAC I IN
        JMP X+2
PUT1,   LAC I IN
        XOR CAC
        AND (777000)
        XOR CAC
        DAC I IN
        JMP X+1

GETN,   SETOUT
GET,    LAC OUT
GETA,   DAC CAC
GETAC,  RAL
        LAC I CAC
        SNL
        LRS 11
        AND (777)
        JMP X

SETOUT, MPTR
        DAC OUT
        JMP X
MPTR,   RCR
        SZL
        XOR (400000)
        ADD (BUF)
        JMP X
SETIN,  MPTR
        JMP SETIN1
PUTI,   PUT
INCIN,  LAC (400000)
INCI1,  ADD IN
SETIN1, DAC IN
        JMP X+1
DECIN,  LAC (377777)
        JMP INCI1
IN,     0
OUT,    0
```

```
CRLF,   LAW 15
TYO,    AND (177)
        SAD (177)
        JMP X+1
        SAD (175)
        LAW 44
        SAD (11)
        JMP TTAB
TYO1,   TSF
        JMP .-1
        TLS 40
        SAD (15)
        JMP TCR
        SAD (12)
        JMP X+1
        ISZ TBC
        JMP X+1
TBC,    0
TTAB,   LAC CAC
        PUSH
        LACQ
        PUSH
        LAW 40
        TYO
        LAC TBC
        IDIVS I
        5
        SZA
        JMP TTAB+4
        POP
        LMQ
        POP
        JMP X
TCR,    LAC (12)
        DZM TBC
        JMP TYO1

TYI,    KSF
        JMP .-1
        KRB
        AND (177)
        DAC CAC
        JMP TYO
```

```
MIN,    LAC (MII)
        JMP NEWOP
MII,    CMA
        ADD SYL
        CMA
        JMP X

INS,    DZM CHRCT
        LAC CPTR
        DAC CPTRS
INS2,   RCH
        SAD (175)
        JMP INS1
        ISZ CHRCT
        JMP INS2
INS1,   LAC CHRCT
        SNA
        JMP X
        NROOM
        SETIN
        LAC CPTRS
        DAC OUT
        LAC CHRCT
        CMA
        TAD (1)
        DAC CNTR
        JMP COPYF
TYPE,   LAW TYO
OUTX,   DZM CNTR
        AND (I-1)
        DAC TYPEM
TYPE1,  LAC CNTR
        SAD Z
        JMP X
        SAD PT
        JMP TYPE2
TYPE3,  MPTR
        GETAC
TYPEM,  TYO
        ISZ CNTR
        KSF
        JMP TYPE1
        KRB
        AND (177)
        SAD (40)
        JMP BEG
        JMP TYPE1

TYPE2,  LAM 134-TYO
        ADD TYPEM
        SAD (134)
        TYO
        LAC PT
        JMP TYPE3
ADDZ,   ADD Z
        SAD (-0)
        CLA
        DAC Z
        JMP X

NROOM,  LAC Z
        DAC OLDZ
        LAC PT
        DAC OLDPT
        ADD CAC
        SAD (-0)
        CLA
        DAC PT
        LAC Z
        CMA
        TAD OLDPT
        TAD (1)
        SMA
        JMP NROOMY
        DAC CNTR
        LAC CAC
        SPA
        JMP NROOMD
        ADDZ
        TAD CLAM
        SETIN
CLAM,   LAM
        TAD OLDZ
        SETOUT
        COPYB
NROOMX, LAC OLDPT
        JMP X
NROOMD, ADDZ
        LAC OLDPT
        SETOUT
        LAC PT
        SETIN
        JMP COPYF
```

```
NROOMY, LAC CAC
        ADDZ
        JMP NROOMX
COPYB,  LAC (377777)
        JMP COPYC
COPYF,  LAC (400000)
COPYC,  DAC INCR
NROOML, GET
        PUT
        LAC INCR
        ADD IN
        DAC IN
        LAC INCR
        ADD OUT
        DAC OUT
        ISZ CNTR
        JMP NROOML
        JMP X

FBL,    LAC PT
FBL1,   DAC CNTR
        SNA
        JMP X
        GETN
        SAD (15)
        JMP FBL2
        LAM
        TAD CNTR
        JMP FBL1

FBL2,   ISZ CNTR
        LAC CNTR
        JMP X

BLIN,   FBL
        JMP SETPT

TLIN,   FBL
        LAW TYO
        JMP OUTX+1

DEL1,   DAC DEL1
        SPA!CMA
        JMP DEL2
        REVER
        LAC DEL1
        CMA
        JMP DEL3
DEL2,   TAD PT
        TAD (1)
        SPA
        ERR
        LAC DEL1
        DAC CAC
        JMP NROOM

REVER,  CMA
CHARS,  ADD PT
SETPT,  SAD (-0)
        CLA
        DAC CAC
        SPA!CMA
        ERR
        ADD Z
        SPA
        SAD (-0)
        JMP POK
        ERR
POK,    LAC CAC
        DAC PT
        JMP X
JUMP,   SNL
        CLA
        JMP SETPT

DPT,    SMA
        JMP DPT1
        PUSH
        LAW 55
        TYO
        POP
        CMA
DPT1,   IDIVS I
        12
        PUSH
        LACQ
        SZA
        DPT1
        POP
        TAD (60)
        JMP TYO

NPRNT,  DPT
        JMP CRLF

ERR,    LAW 77
        TYO
        JMP BEG
```

```
VT,     DAC VTE
        LAW VTB-1
GT,     DAC 10
        LAC CAC
        SAD I 10
        JMP X
        JMP .-2
LMRT,   DAC LMTE
        LAW LMT-1
        JMP GT

VTB,    101     105     111     117     125     131
VTE,    0
LMT,    114     115     122
LMTE,   0

CLUST,  LAW CLT-1
        DAC 11
        LAC IN
        DAC OLDZ
        LAC (SZA
        DAC CLTM
CL1,    LAC OLDZ
        GTLETN
        JMP CL3
        VT
        LAW VTE
        XOR 10
CLTM,   XX
        JMP CL2
CL1A,   ISZ OLDZ
        JMP CL1
CL2,    LAW CLT3
        SAD 11
        JMP CL1A
        CLS
        LAC (1000)
        XOR CLTM
        DAC CLTM
        JMP CL1
CL3,    ISZ OLDZ
CLS,    LAC OLDZ
        DAC I 11
        JMP X
CLT1,   0
CLT2,   0
CLT3,   0
CLT4,   0
```

```
VCL,    CLUST
        LAW CLT4
        SAD I1
        JMP VC1
ITS1,   LAC (210000)
        JMP X
VC1,    LAW ENDT-1
        DAC I1
        LAC CLT3
        DAC IN
VC2,    LAC I I1
        SPA
        JMP ITS2
        TAD (ENTT ENTT-BUF-BUF)
        C2W
        SZA
        JMP VC2
        LAC CLT2
        CMA
        TAD CLT3
        SNA
        JMP ITS1
        SAD (1)
        JMP VC3
VC4,    LAM
        TAD CLT3
        GETN
        LMRT
        LAW LMTE
        SAD I0
        JMP ITS1
ITS2,   LAC (220000)
        JMP X
VC3,    LAC CLT2
        GETN
        DAC TARG
        LAC CLT2
        TAD (1)
        GETN
        SAD TARG
        JMP ITS1
        JMP VC4
BUF,    BUF+4000/
ENTT,   105000
        105123
        000105
        104000
        111116
        107000
ENDT,   0       2       5       10      -0
```

```
    DIG=JMP DIG
    PLUS=SAD PLUS
    MIN=SAD MIN
    END,                LAC Z
                DIG+6
    LOC,        LAC PT
                DIG+6
    PT,         0
    OLDPT,      0
    Z,          0
    OLDZ,       0
    CNTR,       0
    CHRCT,      0
    CPTRS,      0
    DELI,       0
    INCR,       0
    DTB,        PLUS
    0           0           0           0           0           0           0
                PLUS
                0
                MIN
JMP LOC
0
DIG             DIG             DIG             DIG             DIG
DIG             DIG             DIG             DIG             DIG
0               0               0
                XCT NPRNT
0               XCT 16000       0

XCT APPEND
0
XCT CHARS
XCT DELE
XCT READ
XCT PROC
0
0
XCT INS
XCT JUMP
0
XCT BLIN
0
0
0
XCT TYPE
0
XCT REVER
XCT SEAR
XCT TLIN
XCT PINCH

0
0
0
0
JMP END
DTB+100/
START BEG
```

Section 2

SYMBOLS IN ALPHABETIC ORDER

ADDZ	4316	END	10667
ALT	2773	ENDT	10662
ALTF	3066	ENT	400
APPEND	261	ENTT	10654
ARG	3070	ENTI	406
ATAB	2737	ERR	4512
BEG	2744	ES	515
BLIN	4423	ESF	534
BUF	4654	ESL	520
CAC	6	FBL	4406
CBUF	3145	FBL1	4407
CD	3000	FBL2	4420
CDG	3016	FEED	242
CDR	3005	FEED1	244
CDX	3045	FQ	2601
CDY	3057	FTF	2546
CHARS	4447	GET	4125
CHIGH	2632	GETA	4126
CHRCT	10700	GETAC	4127
CIA	160	GETN	4124
CLAM	4350	GT	4517
CLASS	2545	GTLETN	412
CLI	4576	HASHAD	2536
CLTM	4556	HASHT	536
CLT2	4577	HIGH	2544
CLT3	4600	IN	4156
CLT4	4601	INCIN	4150
CLUST	4542	INCII	4151
CL1	4550	INCR	10703
CL1A	4560	INS	4240
CL2	4562	INS1	4250
CL3	4572	INS2	4243
CL5	4573	INI	2537
CNTR	10677	ITS1	4606
COMEND	3067	ITS2	4641
COPYB	4367	JUMP	4465
COPYC	4372	LAST	2541
COPYF	4371	LIS	2746
CPTR	3065	LI1	2752
CPTRS	10701	LI2	2754
CRLF	4160	LMRT	4524
C2W	463	LMT	4536
C2WL	466	LMTE	4541
C2WLOW	513	LOC	10671
C2WN	465	LOW	2542
C2WI	505	MATN	314
DECIN	4154	MATI	317
DELE	4430	MAT2	320
DEL1	10702	MAT3	325
DEL2	4437	MIN	544232
DIG	603073	MII	4234
DPT	4470	MPTR	4140
DPTI	4477	NEWOP	3003
DTB	10704	NPRNT	4510

NROOM	4323
NROOMD	4356
NROOML	4373
NROOMX	4354
NROOMY	4364
NU	2630
NUM	3071
NUMTAG	452
OLDPT	10674
OLDZ	10676
OLDZS	411
OP	3034
OP1	3033
OP2	3043
ORD	378
OUT	4157
OUTX	4265
P	44
POL	45
PINCH	225
PLUS	543002
POK	4462
POP	100151
PPA	250
PPAI	252
PRDAT	2662
PRD1	2674
PRD2	2666
PROC	330
PROUT	2547
PRSIX	2735
PRI	2553
PRIA	2567
PR2	2576
PR3	2603
PR4	2636
PR5	2615
PR6	2623
PR7	2641
PR8	2651
PR9	2644
PT	10673
PUSH	100145
PUT	4105
PUTI	4147
PUTI	4116
RCH	3047
READ	257
READI	263
REVER	4446
RPA	275
RPA1	307
RPA2	277
RPA3	302
RUB	2763
RWD	344
RWL	350
RZ	32
SFAR	163
SEG	167
SEM1	212
SETIN	4145
SETIN1	4158
SETOUT	4135
SETPT	4450
SE1	171
SE2	175
SE3	204
SE4	201
SFD	216
SNFD	223
SPT	236
SYL	3072
TAG	2540
TAGA	456
TAGOK	457
TANG	2543
TBC	4201
TCR	4221
TLEIN	413
TLIN	4425
TSP	2722
TSPA	2724
TTAB	4202
TTAG	432
TYI	4224
TYO	4161
TYO1	4170
TYPE	4264
TYPEM	4277
TYPEI	4270
TYPE2	4310
TYPE3	4275
TYW	2725
TYWI	2730
T3S	2723
VC1	4610
VC2	4614
VC3	4643
VC4	4632
VOCL	4602
VT	4515
VTB	4527
VTE	4535
WRDE	365
X	34
XI	33
YESLN	430
Z	10675

Section 3

SYMBOLS IN NUMERIC ORDER

CAC	6
RZ	32
X1	33
X	34
P	44
POL	45
CIA	160
SEAR	163
SEG	167
SE1	171
SE2	175
SE4	201
SE3	204
SEM1	212
SFD	216
SNFD	223
PINCH	225
FEED	242
FEED1	244
PPA	250
PPA1	252
SPT	256
READ	257
APPEND	261
READ1	263
RPA	275
RPA2	277
RPA3	302
RPA1	307
MATN	314
MAT1	317
MAT2	320
MAT3	325
PROC	330
RWD	344
RWL	350
WRDE	365
ORD	372
ENT	400
ENT1	406
OLDZS	411
GTLETN	412
ILETN	413
YESLN	430
TTAG	432
NUMTAG	452
TAGA	456
TAGOK	457
C2W	463
C2WN	465
C2WL	466
C2W1	505
C2WLOW	513
ES	515
ESL	520
ESF	534
HASHT	536
HASHAD	2536
IN1	2537
TAG	2540
LAST	2541
LOW	2542
TARG	2543
HIGH	2544
CLASS	2545
FTF	2546
PROUT	2547
PR1	2553
PR1A	2567
PR2	2576
PR3	2603
PR5	2615
PR6	2623
NU	2630
FQ	2631
CHIGH	2632
PR4	2636
PR7	2641
PR9	2644
PR8	2651
PRDAT	2662
PRD2	2666
PRD1	2674
TSP	2722
T3S	2723
TSPA	2724
TYW	2725
TYW1	2730
PRSIX	2735
ATAB	2737
BEG	2744
LIS	2746
LI1	2752
LI2	2754
RUB	2763
ALT	2773
CD	3000
NEWOP	3003
CDR	3005
CDG	3016
OP1	3033
OP	3034
OP2	3043
CDX	3045
RCH	3047
CDY	3057
CPTR	3065

ALTF	3066
COMEND	3067
ARG	3070
NUM	3071
SYL	3072
CBUF	3105
PUT	4105
PUTI	4116
GETN	4124
GET	4125
GETA	4126
GETAC	4127
SETOUT	4135
MPTR	4140
SETIN	4145
PUTI	4147
INCIN	4150
INCII	4151
SETINI	4152
DECIN	4154
IN	4156
OUT	4157
CRLF	4160
TYO	4161
TYOI	4170
TBC	4201
TTAB	4202
TCR	4221
TYI	4224
MII	4234
INS	4240
INS2	4243
INSI	4250
TYPE	4264
OUTX	4265
TYPEI	4270
TYPE3	4275
TYPEM	4277
TYPE2	4310
ADDZ	4316
NROOM	4323
CLAM	4350
NROOMX	4354
NROOMD	4356
NROOMY	4364
COPYB	4367
COPYF	4371
COPYC	4372
NROOML	4373
FBL	4406
FBLI	4407
FBL2	4420
BLIN	4423
ILIN	4425
DELE	4430
DEL2	4437
NEVER	4446
CHARS	4447
SETPT	4450
POK	4462
JUMP	4465
OPT	4470
OPTI	4477
NPRNT	4510
ERR	4512
VT	4515
BT	4517
LMRT	4524
VTB	4527
VTE	4535
LMT	4536
LMTE	4541
CLUST	4542
CLI	4550
CLTM	4556
CLIA	4560
CL8	4568
CL3	4572
CL5	4573
CLT	4576
CLT2	4577
CLT3	4600
CLT4	4601
WCL	4602
ITS1	4606
VC1	4610
VC2	4614
VC4	4632
ITS2	4641
VC3	4643
BUF	4654
ENTT	10654
ENDT	10662
END	10667
LOC	10671
PT	10673
CLOPT	10674
Z	10675
O.DZ	10676
CNTR	10677
CHRCT	10700
CPTRS	10701
DELI	10702
INCR	10703
DTB	10704
PUSH	100145
POP	100151
PLUS	543002
MIN	544232
DIG	603073

Section 4

MNEMONIC KEY

ADDZ	flush minus zeroes
ALT	alternate mode
ALTF	flag for 2 alternate modes in succession
APPEND	append to buffer
ARG	argument
ATAB	a table with 3, 4, 2, 1 in it
BEG	beginning
BLIN	beginning of line
BUF	buffer
CAC	contents of accumulator (AC)
CBUF	command buffer
CD	command dispatch
CDG	get command and dispatch it
CDR	read another character of command
CDX	execute call to routine
CDY	" , variation
CHARS	advance ... characters
CHIGH	compare with high
CHRCT	character count
CIA	complement and increment the AC
CLAM	constant LAM (load accumulator with minus)
CLASS	class 3, 4, 2, 1
CLT	cluster table
CLTM	cluster, terminate
CLT2	cluster table, variation
CLT3	"
CLT4	"
CLUST	clustering routine
CL1	" , variation
CL1A	"
CL2	"
CL3	"
CL5	"
CNTR	counter
COMEND	command end
COPYB	copy back
COPYC	copy common
COPYF	copy forward
CPTR	command pointer
CPTRS	command pointer, saved
CRLF	carriage return, line feed
C2W	compare two words
C2WL	compare 2 words loop
C2WLOW	compare 2 words, low test
C2WN	compare two words, variation
C2W1	"
DECIN	decrement IN
DELE	delete
DEL1	" , variation
DEL2	"
DIG	digit

DPT	convert binary to decimal, and print it
DPT1	" , variation
DTB	dispatch table
END	end of buffer
ENDT	ending table (table of pointers)
ENT	ending table, variation
ENTT	table containing actual E, ES, ED, ING endings
ENT1	ending table, variation
ERR	error (will type a question mark and go back to listen)
ES	evaluate symbol
ESF	" , find
ESL	" , loop
FBL	find beginning of line
FBL1	" , variation
FBL2	"
FEED	feed
FEED1	" , variation
FQ	frequency
FTF	first time flag
GET	get AC out from buffer
GETA	" , variation
GETAC	"
GETN	get nth character from buffer
GT	general test (works only for letters)
GTLETN	get a char and test for letter or number
HASHAD	hash address
HASHT	hash table
HIGH	high
IN	pointer to beginning of word
INCIN	increment IN
INCI1	" , variation
INCR	increment of moving
INS	insert
INS1	" , variation
INS2	"
IN1	IN pointer, variation
ITS1	result is ...
ITS2	" , variation
JUMP	jump forward or backward by character
LAST	pointer of last word which is tagged
LIS	listen
LI1	" , variation
LI2	"
LMRT	L, M, R test
LMT	L, M, R table
LMTE	end of L, M, R table
LOC	location for execute command
LOW	low
MATN	match N next word
MAT1	" , variation
MAT2	"
MAT3	"
MIN	minus
MI1	minus, variation
MPTR	make pointer
NEWOP	new operation
NPRNT	print a number

NROOM	make N characters of room in buffer
NROOMD	delete N characters from buffer
NROOML	loop for making room
NROOMX	exit from subroutine NROOM
NROOMY	NROOM subroutine, variation
NU	number
NUM	number
NUM TAG	number tag
OLDPT	old pointer
OLDZ	old number of characters in buffer; also, pointer to end of word while figuring out word
OLDZS	", saved
OP	operation, plus or minus
OP1	", variation
OP2	"
ORD	ordinary word, not one character or if one character, not A, I, or O
OUT	OUT subroutine
OUTX	" , exit
P	pointer to push-down-list
PDL	push-down-list
PINCH	punch
PLUS	plus
POK	pointer OK
POP	pop push-down-list
PPA	punch paper tape, alphanumeric
PPA1	" , variation
PRDAT	print data
PRD1	" , variation
PRD2	"
PROC	processing
PROUT	print out
PRSIX	print six bit code, the code less 4
PR1	print, variation
PR1A	"
PR2	"
PR3	"
PR4	"
PR5	"
PR6	"
PR7	"
PR8	"
PR9	"
PT	current pointer to buffer
PUSH	push push-down-list
PUT	stores character in left half of machine word if IN is positive
PUTI	stores character in right half of machine word if IN is negative
PUT1	put, and increment IN to the next position
RCH	read character
READ	read
READ1	" , variation
REVER	reverse, move pointer back N characters
RPA	read punch, alphanumeric
RPA1	" , variation
RPA2	"
RPA3	"
RUB	rubout

RWD	read word
RWL	read word, loop
RZ	?
SEAR	search
SEG	?
SEM1	modified location during search for # feature
SETIN	sets a register called IN to point to Nth character in buffer
SETIN1	" , variation
SETOUT	sets a register named OUT to point to Nth character in buffer
SETPT	set pointer
SE1	" , variation
SE2	"
SE3	"
SE4	"
SFD	search, found
SNFD	search, not found
SPT	saved pointer
SYL	SYMBOL (refers to current number typed in command buffer)
TAG	register that holds the tag (3, 4, 2, 1, ...)
TAGA	" , variation
TAGOK	tag OK
TARG	temporary argument
TBC	tab counter
TCR	type a carriage return
TLEIN	test letter N
TLIN	type from beginning of line
TSP	type a space
TSPA	" , variation
TTAB	type a tab
TTAG	test tag
TYI	type in
TYO	type out
TYO1	" , variation
TYPE	type
TYPEM	type modified location (typing or punching)
TYPE1	" , variation
TYPE2	"
TYPE3	"
TYW	type word
TYW1	", variation
T3S	type 3 spaces
VC1	vocabulary class 1
VC2	vocabulary class 2
VC3	vocabulary class 3
VC4	vocabulary class 4
VOCL	vocabulary class
VT	vowel table
VTB	vowel table, variation
VTE	vowel table end
WRDE	word end
X	exit
X1	" , variation
YESLN	yes, a letter or a number
Z	number of characters in buffer

Section 5

TYPEWRITER CODES

Following are the seven-bit octal codes for the typewriter symbols in the program. These are according to the system called ASCII (American Standard Code for Information Interchange).

[space]	4Ø	Ø	1ØØ	BELL	7
!	41	A	1Ø1	TAB	11
"	42	B	1Ø2	L.F.	12
#	43	C	1Ø3	C.R.	15
$	44	D	1Ø4	ALTMODE	175
%	45	E	1Ø5	RUBOUT	177
&	46	F	1Ø6		
'	47	G	1Ø7		
(	5Ø	H	11Ø		
)	51	I	111		
*	52	J	112		
+	53	K	113		
,	54	L	114		
-	55	M	115		
.	56	N	116		
/	57	O	117		
Ø	6Ø	P	12Ø		
1	61	Q	121		
2	62	R	122		
3	63	S	123		
4	64	T	124		
5	65	U	125		
6	66	V	126		
7	67	W	127		
8	7Ø	X	13Ø		
9	71	Y	131		
:	72	Z	132		
;	73	[	133		
<	74	\	134		
=	75	]	135		
>	76	↑	136		
?	77	←	137		

Appendix 2

Eight Sample Explanations and Their Assessment

Outline

1.	Resistant Sporangia	R. W. Emerson and C. M. Wilson
2.	Problems Suitable for Mechanical Integration	D. R. Hartree
3.	Nautical Terms	"The Bluejackets' Manual", United States Naval Inst.
4.	Sampling Principles	W. G. Cochran, F. Mosteller, and J. W. Tukey
5.	Sampling Principles -- Revision	E. C. Berkeley
6.	Screening out Potential Troublemakers	"Principles and Problems of Naval Leadership", by Bureau of Naval Personnel
7.	Screening out Potential Troublemakers -- Revision	E. C. Berkeley
8.	Communication is Subtle	The New Yorker

This appendix shows how the computer program described in Appendix 1 operates on a number of different sample explanations, to give the inventory and frequency of their vocabularies in each of:

Class 3, the Special Vocabulary;

Class 4, the Allusion Vocabulary;

Class 2, the Key Vocabulary;

Class 1, the One-Syllable Vocabulary.

For each sample explanation, there is stated:

- the explanation as written in ordinary fashion;
- the explanation as it is put into the computer program (after all corrections and emendations);
- the number and frequency of the vocabulary in each of the four explanation classes, as it is output by the computer program.

1. Resistant Sporangia

The following is part of the third paragraph, from "The Significance of Meiosis in Allomyces" by Ralph W. Emerson and Charles M. Wilson in Science, Vol. 110, July 22, 1949, pp. 86-88.

Resistant sporangia formed by sporophytic thalli grown on slants of yeast-starch agar ordinarily become capable of germination three to six weeks after their formation. At this time each sporangium contains about a dozen expanded, diploid nuclei in an advanced prophase stage. These sporangia are fully mature and, if air dried, they will remain viable and their nuclei will remain in prophase without any further detectable change for periods up to at least ten years. When mature resistant sporangia are taken directly from moist agar cultures and placed in water at 20^{0} to 25^{0} C., they form and release spores in 100 to 130 minutes. During this short interval the two meiotic nuclear divisions occur, and are immediately followed by cleavage of the cytoplasm and organization of the zoospores. Each of these zoospores is haploid and normally uninucleate.

RESISTANT 3 SPORANGIA 3 FORMED BY SPOROPHYTIC 3 THALLI 3 GROWN ON SLANTS 3 OF YEAST-STARCH 3 AGAR 3 ORDINARILY BECOME CAPABLE OF GERMINATION 3 THREE TO SIX WEEKS AFTER THEIR FORMATION. AT THIS TIME EACH SPORANGIUM 3 CONTAINS ABOUT A DOZEN EXPANDED 3, DIPLOID 3 NUCLEI 3 IN AN ADVANCED PROPHASE 3 STAGE. THESE SPORANGIA ARE FULLY MATURE AND, IF AIR-DRIED, WILL REMAIN VIABLE AND THEIR NUCLEI WILL REMAIN IN PROPHASE WITHOUT ANY FURTHER DETECTABLE CHANGE FOR PERIODS UP TO AT LEAST TEN YEARS. WHEN MATURE RESISTANT SPORANGIA ARE TAKEN DIRECTLY FROM MOIST AGAR CULTURES 3 AND PLACED IN WATER AT 20 DEGREES TO 25 DEGREES CENTIGRADE 3, THEY FORM AND RELEASE SPORES 3 IN 100 TO 130 MINUTES. DURING 2\ THIS SHORT INTERVAL THE TWO MEIOTIC 3 NUCLEAR 3 DIVISIONS 3 OCCUR, AND ARE IMMEDIATELY FOLLOWED BY CLEAVAGE 3 OF THE CYTOPLASM 3 AND ORGANIZATION OF THE ZOOSPORES 3. EACH OF THESE ZOOSPORES IS HAPLOID 3 AND NORMALLY UNINUCLEATE 3.

CLASS 3

AGAR	2
CENTIGRADE	1
CLEAVAGE	1
CULTURES	1
CYTOPLASM	1
DIPLOID	1
DIVISIONS	1
EXPANDED	1
GERMINATION	1
HAPLOID	1
MEIOTIC	1
NUCLEAR	1
NUCLEI	2
PROPHASE	2
RESISTANT	2
SLANTS	1
SPORANGIA	3
SPORANGIUM	1
SPORES	1
SPOROPHYTIC	1
THALLI	1
UNINUCLEATE	1
YEAST-STARCH	1
ZOOSPORES	2

NO 24 FQ 31

CLASS 2

ABOUT	1
ADVANCED	1
AFTER	1
AIR-DRIED	1
ANY	1
BECOME	1
CAPABLE	1
CONTAINS	1
DEGREES	2
DETECTABLE	1
DIRECTLY	1
DOZEN	1
DURING	1
FOLLOWED	1
FORMATION	1
FORMED	1
FULLY	1
FURTHER	1
IMMEDIATELY	1
INTERVAL	1
MATURE	2
MINUTES	1
NORMALLY	1
OCCUR	1
ORDINARILY	1
ORGANIZATION	1
PERIODS	1
RELEASE	1
REMAIN	2
TAKEN	1
VIABLE	1
WATER	1
WITHOUT	1

NO 33 FQ 36

CLASS 1

100	1
130	1
20	1
25	1
A	1
AN	1
AND	7
ARE	3
AT	3
BY	2
CHANGE	1
EACH	2
FOR	1
FORM	1
FROM	1
GROWN	1
IF	1
IN	4
IS	1
LEAST	1
MOIST	1
OF	5
ON	1
PLACED	1
SHORT	1
SIX	1
STAGE	1
TEN	1
THE	3
THEIR	2
THESE	2
THEY	1
THIS	2
THREE	1
TIME	1
TO	4
TWO	1
UP	1
WEEKS	1
WHEN	1
WILL	2
YEARS	1

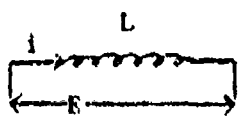
NO 42 FQ 70

2. Problems Suitable for Mechanical Integration

The following is an excerpt from Chapter 2 of "Calculating Instruments and Machines" by Douglas R. Hartree, published by The University of Illinois Press, Urbana, Ill., 1949

Before considering the differential analyser, it will be well to examine the general nature of the situation with which it deals, and what are the characteristic features of problems giving rise to this situation.

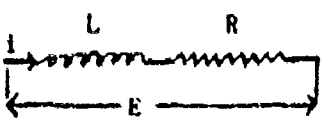
In many applications of mathematics to problems of pure and applied science, there occurs the idea of a rate of change, usually with respect to time or to space, of one or more of the relevant quantities, and then the idea of integration is involved in the process of finding the total change in a time or space interval from the rate of change, which will usually be varying through the interval. In contexts of this nature there are two different kinds of situation which may arise, of which simple examples are provided by the responses of two different circuits to an applied voltage, varying in a known way with time (fig. 2).



$$L \frac{di}{dt} = E(t)$$

$$i = \frac{1}{L} \int E(t)\, dt$$

Fig. 2a



$$L \frac{di}{dt} + Ri = E(t)$$

$$i = \frac{1}{L} \int [E(t) - Ri]\, dt$$

Fig. 2b

If the voltage is applied to a circuit with inductance but with negligible resistance (as in fig. 2a), the time rate of change of current at any moment depends only on the value of the voltage at that moment, which is given and is independent of the value the current itself happens to have. Then we are concerned simply with the evaluation of an integral whose integrand is a known function of the independent variable.

But it very often happens that the rate of change of a quantity at any moment (or point) depends on the magnitude of that quantity itself at that moment (or point). For example, if the voltage is applied to a circuit with inductance and resistance (as in fig. 2b), the rate of change of current depends on the instantaneous value of the current itself, as well as on the voltage. The formal expression of such a situation is what is called in mathematics a <u>differential equation</u>, and the determination of the current at any time involves the solution of this differential equation. This solution can be regarded as the result of evaluating an integral in which the integrand at any time depends in a definite way on the result of the integration up to that time. For example in the circuit illustrated in fig. 2b, the equation for the current can be written

$$i = \frac{1}{L} \int [E(t) - Ri] \, dt,$$

and in the integral here, the current i to be found occurs in one contribution to the integrand.

This aspect of a differential equation is not prominent in the conventional formal treatment of such equations, but it expresses rather closely the way in which it is often best to consider their mechanical solution. Indeed, from the point of view of mechanical integration, it is just this feature which distinguishes the evaluation of a solution of a differential equation from the evaluation of a simple integral of a known function of the independent variable. Thus the essential points in a mechanical instrument for integrating differential equations are an integrating mechanism and means of furnishing to that mechanism as integrand a quantity depending in a definite way on the value of the integral calculated by it.

BEFORE CONSIDERING THE DIFFERENTIAL-ANALYSER 3, IT WILL BE WELL TO EXAMINE THE GENERAL NATURE OF THE SITUATION WITH WHICH IT DEALS, AND WHAT ARE THE CHARACTERISTIC FEATURES OF PROBLEMS GIVING RISE TO THIS SITUATION.

IN MANY APPLICATIONS OF MATHEMATICS TO PROBLEMS IN PURE AND APPLIED SCIENCE, THERE OCCURS THE IDEA OF A RATE-OF-CHANGE 3, USUALLY WITH RESPECT TO TIME OR TO SPACE, OF ONE OR MORE OF THE RELEVANT QUANTITIES , AND THE IDEA OF INTEGRATION 3 IS INVOLVED IN THE PROCESS OF FINDING THE TOTAL CHANGE IN A TIME OR SPACE INTERVAL FROM THE RATE OF CHANGE , WHICH WILL USUALLY BE VARYING THROUGH THE INTERVAL. IN CONTEXTS OF THIS NATURE THERE ARE TWO DIFFERENT KINDS OF SITUATION WHICH MAY ARISE, OF WHICH SIMPLE EXAMPLES ARE PROVIDED BY THE RESPONSES OF TWO DIFFERENT CIRCUITS 3 TO AN APPLIED-VOLTAGE 3, VARYING 3 IN A KNOWN WAY WITH TIME (FIGURE TWO).

IF THE VOLTAGE IS APPLIED TO A CIRCUIT 3 WITH INDUCTANCE 3 BUT WITH NEGLIGIBLE RESISTANCE 3 (AS IN FIGURE 2A), THE TIME-RATE-OF-CHANGE 3 OF CURRENT 3 AT ANY MOMENT DEPENDS ONLY ON THE VALUE 3 OF THE VOLTAGE 3 AT THAT MOMENT, WHICH IS GIVEN AND IS INDEPENDENT 3 OF THE VALUE THE CURRENT ITSELF HAPPENS TO HAVE. THEN WE ARE CONCERNED SIMPLY WITH THE EVALUATION OF AN INTEGRAL 3 WHOSE INTEGRAND 3 IS A KNOWN FUNCTION 3 OF THE INDEPENDENT-VARIABLE 3.

BUT IT VERY OFTEN HAPPENS THAT THE RATE-OF-CHANGE OF A QUANTITY AT ANY MOMENT (OR POINT) DEPENDS ON THE MAGNITUDE OF THAT QUANTITY ITSELF AT THAT MOMENT (OR POINT). FOR EXAMPLE, IF THE VOLTAGE 3 IS APPLIED 3 TO A CIRCUIT 3 WITH INDUCTANCE 3 AND RESISTANCE 3 (AS IN FIGURE 2B), THE RATE OF CHANGE OF CURRENT 3 DEPENDS ON THE INSTANTANEOUS VALUE OF THE CURRENT ITSELF, AS WELL AS ON THE VOLTAGE. THE FORMAL EXPRESSION OF SUCH A SITUATION IS WHAT IS CALLED IN MATHEMATICS A "DIFFERENTIAL-EQUATION 3", AND THE DETERMINATION OF THE CURRENT AT ANY TIME INVOLVES THE SOLUTION OF THIS " DIFFERENTIAL-EQUATION". THIS SOLUTION CAN BE REGARDED AS THE RESULT OF EVALUATING AN INTEGRAL 3 IN WHICH THE INTEGRAND 3 AT ANY TIME DEPENDS IN A DEFINITE WAY ON THE RESULT OF THE INTEGRATION 3 UP TO THAT TIME. FOR EXAMPLE, IN THE CIRCUIT 3 ILLUSTRATED IN FIGURE 2B, THE EQUATION 3 FOR THE CURRENT CAN BE WRITTEN ..... AND IN THE INTEGRAL 3 HERE, THE CURRENT -I- 4 TO BE FOUND OCCURS IN ONE CONTRIBUTION TO THE INTEGRAND.

THIS ASPECT OF A DIFFERENTIAL-EQUATION IS NOT PROMINENT IN THE CONVENTIONAL FORMAL TREATMENT OF SUCH EQUATIONS 3, BUT IT EXPRESSES RATHER CLOSELY THE WAY IN WHICH IT IS OFTEN BEST TO CONSIDER THEIR MECHANICAL SOLUTION. INDEED, FROM THE POINT OF VIEW OF MECHANICAL-INTEGRATION 3, IT IS JUST THIS FEATURE WHICH DISTINGUISHES THE EVALUATION OF A SOLUTION OF A DIFFERENTIAL-EQUATION FROM THE EVALUATION OF A SIMPLE INTEGRAL 3 OF A KNOWN FUNCTION 3 OF THE INDEPENDENT-VARIABLE 3. THUS THE ESSENTIAL POINTS IN A MECHANICAL INSTRUMENT FOR INTEGRATING 3 DIFFERENTIAL-EQUATIONS 3 ARE AN INTEGRATING-MECHANISM 3 AND MEANS OF FURNISHING TO THAT MECHANISM AS INTEGRAND 3 A QUANTITY DEPENDING IN A DEFINITE WAY ON THE VALUE 3 OF THE INTEGRAL 3 CALCULATED BY IT.

CLASS 3

APPLIED	3
APPLIED-VOLTAGE	1
CIRCUIT	3
CIRCUITS	1
CURRENT	7
DIFFERENTIAL-ANALYSER	1
DIFFERENTIAL-EQUATION	4
DIFFERENTIAL-EQUATIONS	1
EQUATION	1
EQUATIONS	1
FUNCTION	2
INDEPENDENT	1
INDEPENDENT-VARIABLE	2
INDUCTANCE	2
INTEGRAL	5
INTEGRAND	4
INTEGRATING	1
INTEGRATING-MECHANISM	1
INTEGRATION	2
MECHANICAL-INTEGRATION	1
RATE-OF-CHANGE	2
RESISTANCE	2
TIME-RATE-OF-CHANGE	1
VALUE	4
VARYING	2
VOLTAGE	4

NU 26 FQ 59

CLASS 4

-I-	1

NU 1 FQ 1

CLASS 2

ANY	4
APPLICATIONS	1
ARISE	1
ASPECT	1
BEFORE	1
CALCULATED	1
CHARACTERISTIC	1
CLOSELY	1
CONCERNED	1
CONSIDER	1
CONSIDERING	1
CONTEXTS	1
CONTRIBUTION	1
CONVENTIONAL	1
DEFINITE	2
DEPENDING	1
DEPENDS	4
DETERMINATION	1
DIFFERENT	2
DISTINGUISHES	1
ESSENTIAL	1
EVALUATING	1
EVALUATION	3
EXAMINE	1
EXAMPLE	2
EXAMPLES	1
EXPRESSES	1
EXPRESSION	1
FEATURE	1
FEATURES	1
FIGURE	4
FORMAL	2
FURNISHING	1
GENERAL	1
GIVEN	1
HAPPENS	2
IDEA	2
ILLUSTRATED	1
INDEED	1
INSTANTANEOUS	1
INSTRUMENT	1
INTERVAL	2
INVOLVED	1
INVOLVES	1
ITSELF	3
MAGNITUDE	1
MANY	1
MATHEMATICS	2
MECHANICAL	2
MECHANISM	1
MOMENT	4
NATURE	2
NEGLIGIBLE	1
OCCURS	2
OFTEN	2
ONLY	1
PROBLEMS	2
PROCESS	1
PROMINENT	1
PROVIDED	1
QUANTITIES	1
QUANTITY	3
RATHER	1
REGARDED	1
RELEVANT	1
RESPECT	1
RESPONSES	1
RESULT	2
SIMPLE	2
SIMPLY	1
SITUATION	4
SOLUTION	4
TOTAL	1
TREATMENT	1
USUALLY	2
VERY	1
WRITTEN	1

NU 77 FQ 117

CLASS 1

2A	1	FOR	4	OF	40
2B	2	FOUND	1	ON	6
A	18	FROM	3	ONE	2
AN	4	GIVING	1	OR	5
AND	8	HAVE	1	POINT	3
ARE	5	HERE	1	POINTS	1
AS	6	IF	2	PURE	1
AT	6	IN	18	RATE	2
BE	5	IS	11	RISE	1
BEST	1	IT	7	SCIENCE	1
BUT	3	JUST	1	SPACE	2
BY	2	KINDS	1	SUCH	2
CALLED	1	KNOWN	3	THAT	6
CAN	2	MAY	1	THE	50
CHANGE	3	MEANS	1	THEIR	1
DEALS	1	MORE	1	THEN	1
FINDING	1	NOT	1	THERE	2
				THIS	6
				THROUGH	1
				THUS	1
				TIME	6
				TO	14
				TWO	3
				UP	1
				VIEW	1
				WAY	4
				WE	1
				WELL	2
				WHAT	2
				WHICH	8
				WHOSE	1
				WILL	2
				WITH	7

NU 67 FQ 313

3. Nautical Terms

The following consists of the first six paragraphs of Chapter 12, "Facts about Ships", pp 210-211, in "The Bluejackets' Manual", 17th Edition, revised by J. V. Noel and W. J. Miller, published by the United States Naval Institute, Annapolis, Md., third printing, August 1965, 684 pp.

A man does not get on a ship; he goes aboard. The front of a ship is the bow, the rear end is the stern. When a man stands at the center of the ship and faces the bow, he faces forward; if he turns around, he faces aft. Facing forward, the right side of the ship is the starboard side; the left side is the port side. An imaginary line from bow to stern is the centerline; it runs fore and aft. The length of this line is the length of the ship. The greatest width of the ship is the beam.

An object directly off the side of the ship is abeam. An object or line running directly across the ship, like a passageway or deck beam, is athwartships. A man standing at the center of the ship is amidships. When he faces either side, he faces outboard. His shipmate at the rail who looks back at him is facing inboard. An object over his head is above; an object underneath him is below.

The floors of a ship are decks; the walls are bulkheads; stairs are ladders. There are no halls or corridors in a ship, only passageways. There is no ceiling in a room aboard ship, only the overhead of the compartment. Openings in the outside of the ship are ports, not windows. Openings through bulkheads, for entering or leaving compartments, are called doors; openings in a deck for passing from one level to another are hatches.

In rough weather, sailors do not shut the windows and lock the doors; they close the ports and dog the hatches. A picture is never nailed to the wall; it is secured to the bulkhead. Sailors don't mop the floor; they swab the deck. The smoke from the ship's fire rooms comes out of a stack, not a chimney.

A sailor doesn't get out of bed in the morning and go to work; he hits the deck and turns to. Even if it is part of his job, he will never be requested to run downstairs to the kitchen and turn on the stove; he may get an order to lay below on the double and light off the galley range.

The words defined above are used when applicable in Naval aircraft as well as afloat in ships and boats.

A MAN DOES NOT GET ON A SHIP; HE GOES-ABOARD 3. THE FRONT OF THE SHIP IS THE BOW 3, THE REAR END IS THE STERN 3. WHEN A MAN STANDS AT THE CENTER OF THE SHIP AND FACES THE BOW, HE FACES FORWARD 3; IF HE TURNS AROUND, HE FACES AFT 3. FACING FORWARD, THE RIGHT SIDE OF THE \SHIP IS THE STARBOARD 3 SIDE; THE LEFT SIDE IS THE PORT 3 SIDE. AN IMAGINARY LINE FROM BOW TO STERN IS THE CENTERLINE 3; IT RUNS FORE-AND-AFT 3. THE GREATEST WIDTH OF THE SHIP IS THE BEAM 3.

AN OBJECT DIRECTLY OFF THE SIDE OF THE SHIP IS ABEAM 3. AN OBJECT OR LINE RUNNING DIRECTLY ACROSS THE SHIP, LIKE A PASSAGEWAY OR DECK BEAM, IS ATHWARTSHIPS 3. A MAN STANDING AT THE CENTER OF THE SHIP IS AMIDSHIPS 3. WHEN HE FACES EITHER \SIDE, HE FACES OUTBOARD 3. HIS SHIPMATE AT THE RAIL WHO LOOKS BACK AT HIM IS FACING INBOARD 3. AN OBJECT OVER HIS HEAD IS ABOVE 3; AN OBJECT UNDERNEATH HIM IS BELOW 3.

THE FLOORS OF A SHIP ARE DECKS 3; THE WALLS ARE BULKHEADS 3; STAIRS ARE LADDERS 3. THERE ARE NO HALLS OR CORRIDORS IN A SHIP, ONLY PASSAGEWAYS 3. THERE IS NO CEILING 2 IN A ROOM ABOARD 3 SHIP, ONLY THE OVERHEAD 3 OF THE COMPARTMENT 3. OPENINGS IN THE SIDE OF THE SHIP ARE PORTS 3, NOT WINDOWS. OPENINGS THROUGH BULKHEADS, FOR ENTERING OR LEAVING COMPARTMENTS 3, ARE CALLED DOORS 3; OPENINGS IN A DECK, FOR PASSING FROM ONE LEVEL TO ANOTHER ARE HATCHES 3.

IN ROUGH WEATHER, SAILORS DO NOT SHUT THE WINDOWS AND LOCK THE DOORS; THEY CLOSE-THE-PORTS 3 AND DOG-THE-HATCHES 3. A PICTURE IS NEVER NAILED TO THE WALL; IT IS SECURED 3 TO THE BULKHEAD 3. SAILORS DON'T MOP THE FLOORS; THEY SWAB-THE-DECK 3. THE SMOKE FROM THE SHIP'S FIRE ROOMS COMES OUT OF A STACK 3, NOT A CHIMNEY.

A SAILOR DOES NOT GET OUT OF BED IN THE MORNING 2\ AND GO TO WORK; HE HITS-THE-DECK 3 AND TURNS-TO 3. EVEN IF IT IS PART OF HIS JOB, HE WILL NEVER BE REQUESTED TO RUN DOWN TO THE KITCHEN AND TURN ON THE STOVE; HE MAY GET AN ORDER 3 TO LAY-BELOW 3 ON-THE-DOUBLE 3 AND LIGHT-OFF 3 THE GALLEY-RANGE 3.

THE WORDS DEFINED ABOVE ARE USED WHEN APPLICABLE IN NAVAL AIRCRAFT AS WELL AS AFLOAT IN SHIPS AND BOATS.

CLASS 3

ABOARD	1
BULKHEAD	1
BULKHEADS	2
CLOSE-THE-PORTS	1
COMPARTMENT	1
COMPARTMENTS	1
DECKS	1
DOG-THE-HATCHES	1
DOORS	2
GALLEY-RANGE	1
HATCHES	1
HITS-THE-DECK	1
LADDERS	1
LAY-BELOW	1
LIGHT-OFF	1
ON-THE-DOUBLE	1
ORDER	1
OVERHEAD	1
PASSAGEWAYS	1
PORTS	1
SECURED	1
STACK	1
SWAB-THE-DECK	1
TURNS-TO	1

NU 24 FQ 26

CLASS 2

ABOVE	1
AFLOAT	1
AIRCRAFT	1
ANOTHER	1
APPLICABLE	1
CEILING	1
CHIMNEY	1
CORRIDORS	1
DEFINED	1
ENTERING	1
EVEN	1
KITCHEN	1
LEVEL	1
MORNING	1
NAVAL	1
NEVER	2
ONLY	2
OPENINGS	3
PICTURE	1
REQUESTED	1
SAILOR	1
SAILORS	2
WEATHER	1
WINDOWS	2

NU 24 FQ 30

CLASS 1

A	6
AN	1
AND	7
ARE	8
AS	2
BE	1
BED	1
BOATS	1
CALLED	1
COMES	1
DECK	1
DO	1
DOES	1
DON'T	1
DOWN	1
FIRE	1
FLOOR	1
FLOORS	1
FOR	2
FROM	2
GET	2
GO	1
HALLS	1
HE	3
HIS	1
IF	1
IN	8
IS	4
IT	2
JOB	1
LEAVING	1
LOCK	1
MAY	1
MOP	1
NAILED	1
NO	2
NOT	4
OF	6
ON	1
ONE	1
OR	2
OUT	2
PART	1
PASSING	1
ROOM	1
ROOMS	1
ROUGH	1
RUN	1
SHIP	4
SHIP'S	1
SHIPS	1
SHUT	1
SIDE	1
SMOKE	1
STAIRS	1
STOVE	1
THE	18
THERE	2
THEY	2
THROUGH	1
TO	7
TURN	1
USED	1
WALL	1
WALLS	1
WELL	1
WHEN	1
WILL	1
WORDS	1
WORK	1

NU 70 FQ 146

4. Sampling Principles

The following is taken almost verbatim from the first two paragraphs of "Principles of Sampling", Appendix G, of "Statistical Problems of the Kinsey Report.....: A Report of the American Statistical Association Committee to Advise the National Research Council Committee....." The authors of the committee report are William G. Cochran, Frederick Mosteller, and John W. Tukey. The committee report is published by the American Statistical Association, Washington, D.C., 1954, 338 pages.

Whether by biologists, sociologists, engineers, or chemists, sampling is all too often taken far too lightly. In the early years of the present century, it was not uncommon to measure the claws and carapaces of 1000 crabs, or to count the number of veins in each of 1000 leaves, and then to attach to the results the "probable error" which would have been appropriate had the 1000 crabs or the 1000 leaves been drawn at random from the population of interest. Such actions were unwarranted shot-gun marriages between the quantitatively unsophisticated idea of a sample as "what you get by grabbing a handful" and the mathematically precise notion of a "simple random sample". In the years between we have learned caution by bitter experience. We insist on some semblance of mechanical (dice, coins, random number tables, etc.) randomization before we treat a sample from an existent population as if it were random. We realize that if someone just "grabs a handful", the individuals in the handful almost always resemble one another (on the average) more than do the members of a simple random sample. Even if the "grabs" are randomly spread around so that every individual has an equal chance of entering the sample, there are difficulties. Since the individuals of grab samples resemble one another more than do individuals of random samples, it follows (by a simple mathematical argument) that the means of grab samples resemble one another less than the means of random samples of the same size. For a grab sample therefore we tend to underestimate the variability of the

population, although we should have to overestimate it in order to obtain valid estimates of variability of grab sample means by substituting such an estimate into the formula for the variability of means of simple random samples. Thus using simple random sample formulas for grab sample means introduces double bias, both parts of which lead to an unwarranted appearance of higher stability.

Returning to the crabs, we may suppose that the crabs in which we are interested are all the individuals of a wide-ranging species, spread along a few hundred miles of coast. It is obviously impractical to seek to take a simple random sample from the species — no one knows how to give each crab in the species an equal chance of being drawn into the sample (to say nothing of trying to make these chances independent). But this does not bar us from honestly assessing the likely range of fluctuation of the result.

Much effort has been applied in recent years particularly in sampling human populations to the development of sampling plans which simultaneously

(1) are economically feasible,

(2) give reasonably precise results, and

(3) show within themselves an honest measure of fluctuation of their results.

Any excuse for the dangerous practice of treating non-random samples as random ones is now entirely tenuous. Wide knowledge of the principles involved is needed if scientific investigations involving samples (and what such investigation does not?) are to be solidly based. Additional knowledge of techniques is not so vitally important, though it can lead to substantial economic gain.

WHETHER BY BIOLOGISTS, SOCIOLOGISTS, ENGINEERS, OR CHEMISTS, SAMPLING 3 IS ALL TOO OFTEN TAKEN FAR TOO LIGHTLY. IN THE EARLY YEARS OF THE PRESENT CENTURY, IT WAS NOT UNCOMMON TO MEASURE THE CLAWS AND CARAPACES OF 1000 CRABS, OR TO COUNT THE NUMBER OF VEINS IN EACH OF 1000 LEAVES AND THEN TO ATTACH TO THE RESULTS THE "PROBABLE-ERROR" 3 WHICH WOULD HAVE BEEN APPROPRIATE HAD THE 1000 CRABS OR 1000 LEAVES BEEN DRAWN AT-RANDOM 3 FROM THE POPULATION 3 OF INTEREST. SUCH ACTIONS WERE UNWARRANTED SHOTGUN MARRIAGES BETWEEN THE QUANTITATIVELY UNSOPHISTICATED IDEA OF A SAMPLE 3 AS "WHAT YOU GET BY GRABBING A HANDFUL" AND THE MATHEMATICALLY PRECISE NOTION OF A "SIMPLE-RANDOM-SAMPLE" 3. IN THE YEARS BETWEEN WE HAVE LEARNED CAUTION BY BITTER EXPERIENCE. WE INSIST ON SOME SEMBLANCE OF MECHANICAL-RANDOMIZATION 3 (DICE, COINS, RANDOM-NUMBER-TABLES 3, ETC.) BEFORE WE TREAT A SAMPLE 3 FROM AN EXISTENT-POPULATION 3 AS IF IT WERE RANDOM 3. WE REALIZE THAT IF SOMEONE JUST "GRABS A HANDFUL", THE INDIVIDUALS IN THE HANDFUL ALMOST ALWAYS RESEMBLE ONE ANOTHER (ON-THE-AVERAGE 3) MORE THAN DO THE MEMBERS OF A SIMPLE-RANDOM-SAMPLE 3. EVEN IF THE "GRABS" ARE RANDOMLY 3 SPREAD AROUND SO THAT EVERY INDIVIDUAL HAS AN EQUAL CHANCE OF ENTERING THE SAMPLE 3, THERE ARE DIFFICULTIES. SINCE THE INDIVIDUALS OF GRAB-SAMPLES 3 RESEMBLE ONE ANOTHER MORE THAN DO INDIVIDUALS OF RANDOM-SAMPLES 3, IT FOLLOWS (BY A SIMPLE MATHEMATICAL ARGUMENT) THAT THE MEANS 3 OF GRAB-SAMPLES 3 RESEMBLE ONE ANOTHER LESS THAN THE MEANS 3 OF RANDOM-SAMPLES 3 OF THE SAME SIZE.

FOR A GRAB-SAMPLE 3 THEREFORE WE TEND TO UNDERESTIMATE THE VARIABILITY 3 OF THE POPULATION 3, ALTHOUGH WE SHOULD HAVE TO OVERESTIMATE IT IN ORDER TO OBTAIN VALID 3 ESTIMATES OF VARIABILITY 3 OF GRAB-SAMPLE 3 MEANS 3 BY SUBSTITUTING SUCH AN ESTIMATE INTO THE FORMULA FOR THE VARIABILITY 3 OF MEANS 3 OF SIMPLE-RANDOM-SAMPLES 3. THUS USING SIMPLE-RANDOM-SAMPLES 3 FOR GRAB-SAMPLE 3 MEANS 3 INTRODUCES DOUBLE BIAS 3, BOTH PARTS OF WHICH LEAD TO AN UNWARRANTED APPEARANCE OF HIGHER STABILITY 3.

RETURNING TO THE CRABS, WE MAY SUPPOSE THAT THE CRABS WHICH WE ARE INTERESTED IN ARE ALL THE INDIVIDUALS OF A WIDE-RANGING SPECIES, SPREAD ALONG A FEW HUNDRED MILES OF COAST.

IT IS OBVIOUSLY IMPRACTICAL TO SEEK TO TAKE A SIMPLE-RANDOM-SAMPLE 3 FROM THE SPECIES -- NO ONE KNOWS HOW TO GIVE EACH CRAB IN THE SPECIES AN EQUAL CHANCE OF BEING DRAWN 3 INTO THE SAMPLE 3 (TO SAY NOTHING 2 OF TRYING TO MAKE THESE CHANCES INDEPENDENT 3). BUT THIS DOES NOT BAR US FROM HONESTLY ASSESSING THE LIKELY RANGE 3 OF FLUCTUATION 3 OF THE RESULT.

MUCH EFFORT HAS BEEN APPLIED IN RECENT YEARS, PARTICULARLY IN SAMPLING HUMAN POPULATIONS, TO THE DEVELOPMENT OF SAMPLING-PLANS WHICH SIMULTANEOUSLY ARE ECONOMICALLY FEASIBLE, GIVE REASONABLY PRECISE RESULTS, AND SHOW WITHIN THEMSELVES AN HONEST MEASURE OF FLUCTUATION OF THEIR RESULTS.

ANY EXCUSE FOR THE DANGEROUS PRACTICE OF TREATING NON-RANDOM-SAMPLES AS RANDOM ONES IS NOW ENTIRELY TENUOUS. WIDE KNOWLEDGE OF THE PRINCIPLES INVOLVED IS NEEDED IF SCIENTIFIC INVESTIGATIONS INVOLVING SAMPLES (AND WHAT SUCH INVESTIGATION DOES NOT?) ARE TO BE SOLIDLY BASED. ADDITIONAL KNOWLEDGE OF TECHNIQUES IS NOT SO VITALLY IMPORTANT, THOUGH IT CAN LEAD TO SUBSTANTIAL ECONOMIC GAINS.

```
CLASS 3
AT-RANDOM   1
BIAS        1
DRAWN       2
EXISTENT-POPULATION 1
FLUCTUATION 2
GRAB-SAMPLE 3
GRAB-SAMPLES 2
INDEPENDENT 1
MEANS       5
MECHANICAL-RANDOMIZATION 1
NON-RANDOM-SAMPLES 1
ON-THE-AVERAGE 1
POPULATION 2
POPULATIONS 1
PROBABLE-ERROR 1
RANDOM      2
RANDOM-NUMBER-TABLES 1
RANDOM-SAMPLES 2
RANDOMLY    1
RANGE       1
SAMPLE      4
SAMPLES     1
SAMPLING    2
SAMPLING-PLANS 1
SIMPLE-RANDOM-SAMPLE 3
SIMPLE-RANDOM-SAMPLES 2
STABILITY   1
VALID       1
VARIABILITY 3
NU 29 FQ 50
```

CLASS 2

ACTIONS	1
ADDITIONAL	1
ALMOST	1
ALONG	1
ALTHOUGH	1
ALWAYS	1
ANOTHER	3
ANY	1
APPEARANCE	1
APPLIED	1
APPROPRIATE	1
ARGUMENT	1
AROUND	1
ASSESSING	1
ATTACH	1
BEFORE	1
BETWEEN	2
BIOLOGISTS	1
BITTER	1
CARAPACES	1
CAUTION	1
CENTURY	1
CHEMISTS	1
DANGEROUS	1
DEVELOPMENT	1
DIFFICULTIES	1
DOUBLE	1
EARLY	1
ECONOMIC	1
ECONOMICALLY	1
EFFORT	1
ENGINEERS	1
ENTERING	1
ENTIRELY	1
EQUAL	2
ESTIMATE	1
ESTIMATES	1
EVEN	1
EVERY	1
EXCUSE	1
EXPERIENCE	1
FEASIBLE	1
FOLLOWS	1
FORMULA	1
HANDFUL	3
HIGHER	1
HONEST	1
HONESTLY	1
HUMAN	1
HUNDRED	1
IDEA	1
IMPORTANT	1
IMPRACTICAL	1
INDIVIDUAL	1
INDIVIDUALS	4
INSIST	1
INTEREST	1
INTERESTED	1
INTO	2
INTRODUCES	1
INVESTIGATION	1
INVESTIGATIONS	1
INVOLVED	1
INVOLVING	1
KNOWLEDGE	2
LIGHTLY	1
LIKELY	1
MARRIAGES	1
MATHEMATICAL	1
MATHEMATICALLY	1
MEASURE	2
MEMBERS	1
NOTHING	1
NOTION	1
NUMBER	1
OBTAIN	1
OBVIOUSLY	1
OFTEN	1
ORDER	1
OVERESTIMATE	1
PARTICULARLY	1
PRACTICE	1
PRECISE	2
PRESENT	1
PRINCIPLES	1
QUANTITATIVELY	1
REALIZE	1
REASONABLY	1
RECENT	1
RESEMBLE	3
RESULT	1
RESULTS	3
RETURNING	1
SCIENTIFIC	1
SEMBLANCE	1
SHOTGUN	1
SIMPLE	1
SIMULTANEOUSLY	1
SOCIOLOGISTS	1
SOLIDLY	1
SOMEONE	1
SPECIES	3
SUBSTANTIAL	1
SUBSTITUTING	1
SUPPOSE	1
TAKEN	1
TECHNIQUES	1
TENUOUS	1
THEMSELVES	1
THEREFORE	1
UNCOMMON	1
UNDERESTIMATE	1
UNSOPHISTICATED	1
UNWARRANTED	2
VITALLY	1
WHETHER	1
WIDE-RANGING	1
WITHIN	1

NU 118 FQ 138

CLASS 1

**	1	GRABBING	1	SIZE	1
1000	4	GRABS	2	SO	2
A	11	HAD	1	SOME	1
ALL	2	HAS	2	SPREAD	2
AN	6	HAVE	3	SUCH	3
AND	5	HOW	1	TAKE	1
ARE	6	IF	4	TEND	1
AS	3	IN	9	THAN	3
BAR	1	IS	5	THAT	4
BASED	1	IT	6	THE	35
BE	1	JUST	1	THEIR	1
BEEN	3	KNOWS	1	THEN	1
BEING	1	LEAD	2	THERE	1
BOTH	1	LEARNED	1	THESE	1
BUT	1	LEAVES	2	THIS	1
BY	5	LESS	1	THOUGH	1
CAN	1	MAKE	1	THUS	1
CHANCE	2	MAY	1	TO	17
CHANCES	1	MILES	1	TOO	2
CLAWS	1	MORE	2	TREAT	1
COAST	1	MUCH	1	TREATING	1
COINS	1	NEEDED	1	TRYING	1
COUNT	1	NO	1	US	1
CRAB	1	NOT	4	USING	1
CRABS	4	NOW	1	VEINS	1
DICE	1	OF	34	WAS	1
DO	2	ON	1	WE	8
DOES	2	ONE	4	WERE	2
EACH	2	ONES	1	WHAT	2
ETC	1	OR	3	WHICH	4
FAR	1	PARTS	1	WIDE	1
FEW	1	SAME	1	WOULD	1
FOR	4	SAY	1	YEARS	3
FROM	4	SEEK	1	YOU	1
GAINS	1	SHOULD	1	NU 108 FQ 300	
GET	1	SHOW	1		
GIVE	2	SINCE	1		

5. Sampling Principles — Revision

The following is a revision of the previous explanation.

In scientific investigation, sampling is all too often taken far too lightly. Fifty years ago, it was not uncommon for a scientist to measure the claws and shells of 1000 crabs or to count the number of veins in each of 1000 leaves, and then to attach to his averages a computed degree of variation called the "probable error".

This figure could have been appropriate if the 1000 crabs (or leaves) had been drawn (or selected) from their "population" (the group of all the individuals being studied) "at random", that is, under conditions where each individual has an equal chance of being drawn. But nearly all the time the condition of drawing at random was not observed. Such drawings were unwarranted shotgun marriages between the quantitatively unsophisticated idea of a sample as "what you get by grabbing a handful", and the mathematically precise notion of a "simple random sample", a sample where each individual in the population of interest has an equal chance of being drawn.

Since that time we have learned caution by bitter experience. We insist on at least some approach towards using a mechanical device for selecting at random, a device such as coins or dice or tables of numbers that occur at random, in order to draw individuals for the sample. Only then do we treat a sample from a population that exists (some populations are indefinitely large and do not exist) as if the sample were random.

We realize that if someone just grabs a handful, the individuals in the handful almost always resemble one another (on the average) more than do the individuals in a simple random sample. Even if the grabs are randomly spread around so that every individual has an equal chance of being drawn for the sample,

there are difficulties. The individuals of grab-samples resemble one another more than do the individuals of random samples; so it follows (by a simple mathematical argument) that the means (the averages) of grab-samples of a given size resemble one another less than the means of random samples of that size.

Using grab-samples, therefore, we tend to underestimate the degree of variation of the population. Applying to grab-samples the formulas that are appropriate only for random samples and putting into those formulas the means of grab-samples, introduces bias (error in one direction). This leads to an unwarranted appearance of a smaller degree of variation in the population than actually exists.

Returning to the crabs, we may suppose that the crabs in which we are interested are all the individuals of a wide-ranging species, spread along a few hundred miles of coast. It is obviously impractical to seek to take a simple random sample from the species -- no one knows how to give each crab in the species an equal chance of being drawn into the sample (to say nothing of trying to make these chances independent). But this does not bar us from honestly trying to assess the likely degree of variation in the result.

Much effort has been applied in recent years, particularly in sampling human populations, to develop plans for sampling which simultaneously

(1) are economically feasible,

(2) give reasonably precise results, and

(3) show within themselves an honest degree of variation in their results.

Any excuse for the dangerous practice of treating as random samples samples that are not random is now without a shred of justification. But wide knowledge of the principles of sampling is needed if scientific investigations involving samples (and what such investigation does not?) are to be solidly based. Additional knowledge of techniques is not so vitally important, though it can lead to substantial economic gains.

IN SCIENTIFIC INVESTIGATION, SAMPLING 3 IS ALL TOO OFTEN TAKEN FAR TOO LIGHTLY. FIFTY YEARS AGO, IT WAS NOT UNCOMMON FOR A SCIENTIST TO MEASURE THE CLAWS AND SHELLS OF 1000 CRABS OR TO COUNT THE NUMBER OF VEINS IN EACH OF 1000 LEAVES, AND THEN TO ATTACH TO HIS AVERAGES 3 A COMPUTED DEGREE-OF-VARIATION 3 CALLED THE "PROBABLE-ERROR 3".

THIS FIGURE COULD HAVE BEEN APPROPRIATE IF THE 1000 CRABS (OR LEAVES) HAD BEEN DRAWN 3 (OR SELECTED) FROM THEIR POPULATION 3 (THE GROUP OF ALL INDIVIDUALS BEING STUDIED), "AT-RANDOM 3", THAT IS, UNDER CONDITIONS WHERE EACH INDIVIDUAL HAS AN EQUAL CHANCE OF BEING DRAWN. BUT NEARLY ALL THE TIME THE CONDITION OF DRAWING 3 AT-RANDOM 3 WAS NOT OBSERVED. SUCH DRAWINGS 3 WERE UNWARRANTED SHOTGUN MARRIAGES BETWEEN THE QUANTITATIVELY UNSOPHISTICATED IDEA OF A SAMPLE 3 AS "WHAT YOU GET BY GRABBING A HANDFUL", AND THE MATHEMATICALLY PRECISE NOTION OF A "SIMPLE-RANDOM-SAMPLE", A SAMPLE 3 WHERE EACH INDIVIDUAL IN THE POPULATION 3 OF INTEREST HAS AN EQUAL CHANCE OF BEING DRAWN.

SINCE THAT TIME WE HAVE LEARNED CAUTION BY BITTER EXPERIENCE. WE INSIST ON AT LEAST SOME APPROACH TOWARDS USING MECHANICAL DEVICE FOR SELECTING AT-RANDOM 3, A DEVICE SUCH AS COINS OR DICE OR TABLES OF NUMBERS THAT OCCUR AT-RANDOM 3, IN ORDER TO DRAW 3 INDIVIDUALS FOR THE SAMPLE 3. ONLY THEN DO WE TREAT A SAMPLE 3 FROM A POPULATION 3 THAT EXISTS (SOME POPULATIONS 3 ARE INDEFINITELY LARGE AND DO NOT EXIST) AS IF THE SAMPLE 3 WERE RANDOM 3. WE REALIZE THAT IF SOMEONE JUST GRABS A HANDFUL, THE INDIVIDUALS IN THE HANDFUL ALMOST ALWAYS RESEMBLE ONE ANOTHER (ON-THE-AVERAGE 3) MORE THAN DO THE INDIVIDUALS IN A SIMPLE-RANDOM-SAMPLE 3. EVEN IF THE GRABS ARE RANDOMLY 3 SPREAD AROUND SO THAT EVERY INDIVIDUAL HAS AN EQUAL CHANCE OF BEING DRAWN 3 FOR THE SAMPLE 3, THERE ARE DIFFICULTIES. THE INDIVIDUALS OF GRAB-SAMPLES 3 RESEMBLE ONE ANOTHER MORE THAN DO THE INDIVIDUALS OF RANDOM-SAMPLES 3; SO IT FOLLOWS (BY A SIMPLE MATHEMATICAL ARGUMENT) THAT THE MEANS 3 (THE AVERAGES 3) OF GRAB-SAMPLES 3 OF A GIVEN SIZE RESEMBLE ONE ANOTHER LESS THAN THE MEANS OF RANDOM-SAMPLES OF THAT SIZE.

USING GRAB-SAMPLES 3, THEREFORE, WE TEND TO UNDERESTIMATE THE DEGREE-OF-VARIATION 3 OF THE POPULATION 3. APPLYING TO GRAB-SAMPLES 3 THE FORMULAS THAT ARE APPROPRIATE ONLY FOR RANDOM-SAMPLES 3 AND PUTTING INTO THOSE FORMULAS THE MEANS 3 OF GRAB-SAMPLES 3, INTRODUCES BIAS 3 (ERROR IN ONE DIRECTION). THIS LEADS TO AN UNWARRANTED APPEARANCE OF A SMALLER DEGREE-OF-VARIATION IN THE POPULATION THAN ACTUALLY EXISTS. RETURNING TO THE CRABS, WE MAY SUPPOSE THAT THE CRABS IN WHICH WE ARE INTERESTED ARE ALL THE INDIVIDUALS OF A WIDE-RANGING SPECIES, SPREAD ALONG A FEW HUNDRED MILES OF COAST. IT IS OBVIOUSLY IMPRACTICAL TO SEEK TO TAKE A SIMPLE-RANDOM-SAMPLE 3 FROM THE SPECIES -- NO ONE KNOWS HOW TO GIVE EACH CRAB IN THE SPECIES AN EQUAL CHANCE OF BEING DRAWN 3 INTO THE SAMPLE 3 (TO SAY NOTHING 2 OF TRYING TO MAKE THESE CHANCES INDEPENDENT 3). BUT THIS DOES NOT BAR US FROM HONESTLY ASSESSING THE LIKELY DEGREE-OF-VARIATION 3 IN THE RESULT.

MUCH EFFORT HAS BEEN APPLIED IN RECENT YEARS, PARTICULARLY IN SAMPLING 3 HUMAN POPULATIONS 3, TO DEVELOP PLANS FOR SAMPLING WHICH SIMULTANEOUSLY ARE ECONOMICALLY FEASIBLE, GIVE REASONABLY PRECISE RESULTS, AND SHOW WITHIN THEMSELVES AN HONEST DEGREE-OF-VARIATION 3 IN THEIR RESULTS.

ANY EXCUSE FOR THE DANGEROUS PRACTICE OF TREATING AS RANDOM-SAMPLES 3 SAMPLES 3 THAT ARE NOT RANDOM 3 IS NOW WITHOUT A SHRED OF JUSTIFICATION. BUT WIDE KNOWLEDGE OF THE PRINCIPLES OF SAMPLING 3 IS NEEDED IF SCIENTIFIC INVESTIGATIONS INVOLVING SAMPLES 3 (AND WHAT SUCH INVESTIGATION DOES NOT?) ARE TO BE SOLIDLY BASED. ADDITIONAL KNOWLEDGE OF TECHNIQUES IS NOT SO VITALLY IMPORTANT, THOUGH IT CAN LEAD TO SUBSTANTIAL ECONOMIC GAINS.

CLASS 3
AT-RANDOM 4
AVERAGES 2
BIAS 1
DEGREE-OF-VARIATION 5
DRAW 1
DRAWING 1
DRAWINGS 1
DRAWN 5
GRAB-SAMPLES 5
INDEPENDENT 1
MEANS 3
ON-THE-AVERAGE 1
POPULATION 5
POPULATIONS 2
PROBABLE-ERROR 1
RANDOM 2
RANDOM-SAMPLES 4
RANDOMLY 1
SAMPLE 7
SAMPLES 2
SAMPLING 4
SIMPLE-RANDOM-SAMPLE 3
NU 22 FQ 61

CLASS 2
ACTUALLY 1
ADDITIONAL 1
AGO 1
ALMOST 1
ALONG 1
ALWAYS 1
ANOTHER 3
ANY 1
APPEARANCE 1
APPLIED 1
APPLYING 1
APPROACH 1
APPROPRIATE 2
ARGUMENT 1
AROUND 1
ASSESSING 1
ATTACH 1
BETWEEN 1
BITTER 1
CAUTION 1
COMPUTED 1
CONDITION 1
CONDITIONS 1
DANGEROUS 1
DEVELOP 1
DEVICE 2
DIFFICULTIES 1
DIRECTION 1
ECONOMIC 1
ECONOMICALLY 1
EFFORT 1
EQUAL 4
ERROR 1
EVEN 1
EVERY 1
EXCUSE 1
EXIST 1
EXISTS 2
EXPERIENCE 1
FEASIBLE 1
FIFTY 1
FIGURE 1
FOLLOWS 1
FORMULAS 2
GIVEN 1
HANDFUL 3
HONEST 1
HONESTLY 1
HUMAN 1
HUNDRED 1
IDEA 1
IMPORTANT 1
IMPRACTICAL 1
INDEFINITELY 1
INDIVIDUAL 3
INDIVIDUALS 7
INSIST 1
INTEREST 1
INTERESTED 1
INTO 2
INTRODUCES 1
INVESTIGATION 2
INVESTIGATIONS 1
INVOLVING 1
JUSTIFICATION 1
KNOWLEDGE 2
LIGHTLY 1
LIKELY 1
MARRIAGES 1
MATHEMATICAL 1
MATHEMATICALLY 1
MEASURE 1
MECHANICAL 1
NEARLY 1
NOTHING 1
NOTION 1
NUMBER 1
NUMBERS 1
OBSERVED 1
OBVIOUSLY 1
OCCUR 1
OFTEN 1
ONLY 2
ORDER 1
PARTICULARLY 1
PRACTICE 1
PRECISE 2
PRINCIPLES 1
QUANTITATIVELY 1
REALIZE 1
REASONABLY 1
RECENT 1
RESEMBLE 3
RESULT 1
RESULTS 2
RETURNING 1
SCIENTIFIC 2
SCIENTIST 1
SELECTED 1
SELECTING 1
SHOTGUN 1
SIMPLE 1
SIMULTANEOUSLY 1
SMALLER 1
SOLIDLY 1
SOMEONE 1
SPECIES 3
STUDIED 1
SUBSTANTIAL 1
SUPPOSE 1
TABLES 1
TAKEN 1
TECHNIQUES 1
THEMSELVES 1
THEREFORE 1
TOWARDS 1
UNCOMMON 1
UNDER 1
UNDERESTIMATE 1
UNSOPHISTICATED 1
UNWARRANTED 2
VITALLY 1
WIDE-RANGING 1
WITHIN 1
WITHOUT 1
NU 125 FU 156

CLASS 1

--	1
1000	3
A	19
ALL	4
AN	6
AND	7
ARE	9
AS	4
AT	1
BAR	1
BASED	1
BE	1
BEEN	3
BEING	5
BUT	3
BY	3
CALLED	1
CAN	1
CHANCE	4
CHANCES	1
CLAWS	1
COAST	1
COINS	1
COULD	1
COUNT	1
CRAB	1
CRABS	4
DICE	1
DO	4
DOES	2
EACH	4
FAR	1
FEW	1
FOR	7
FROM	4
GAINS	1
GET	1
GIVE	2
GRABBING	1
GRABS	2
GROUP	1
HAD	1
HAS	4
HAVE	2
HIS	1
HOW	1
IF	5
IN	14
IS	6
IT	4
JUST	1
KNOWS	1
LARGE	1
LEAD	1
LEADS	1
LEARNED	1
LEAST	1
LEAVES	2
LESS	1
MAKE	1
MAY	1
MILES	1
MORE	2
MUCH	1
NEEDED	1
NO	1
NOT	7
NOW	1
OF	30
ON	1
ONE	5
OR	5
PLANS	1
PUTTING	1
SAY	1
SEEK	1
SHELLS	1
SHOW	1
SHRED	1
SINCE	1
SIZE	2
SO	3
SOME	2
SPREAD	2
SUCH	3
TAKE	1
TEND	1
THAN	4
THAT	11
THE	37
THEIR	2
THEN	2
THERE	1
THESE	1
THIS	3
THOSE	1
THOUGH	1
TIME	2
TO	17
TOO	2
TREAT	1
TREATING	1
TRYING	1
US	1
USING	2
VEINS	1
WAS	2
WE	7
WERE	2
WHAT	2
WHERE	2
WHICH	2
WIDE	1
YEARS	2
YOU	1

NU 115 FQ 359

6. Screening out Potential Troublemakers

The following is taken from pp 40-41 of Chapter 8, "The Case of the Suspended Sentence", in "Principles and Problems of Naval Leadership", NAVPERS 15924A, second edition, prepared by the Bureau of Naval Personnel, Washington, D.C., 1964, 107 pp.

If the behavior of human beings were wholly determined by heredity and environment, or by purely mechanical or automatic laws, prediction of future success or failure would be easy. If we could say that human beings, when stimulated in a particular fashion, would automatically react in a given way, not only the predictability, but the control of human behavior would likewise be easy. Indeed, some schools of thought and teaching have given this impression. One such school might loosely be termed "behavioristic," another "mechanistic". Developments in the science of cybernetics (a comparative study of complex mechanical-electronic communications systems, such as computers, with the control system formed by the nervous system and the brain) have led some people to believe that there is no essential difference between men and highly refined "computers".

But the fact of the matter is that no machine ever invented, past, present, or future, no matter how complex its operations, has had or can have a built-in personal moral responsibility. A man can know all the rules. He can study behavioral principles all his life. In turn he can be studied in exhausting detail. His blood can be tested; x-rays and fluoroscopes can strip him of every physiological secret. Electroencephalograms, psychogalvanometers, and Rorschach diagnoses can tell a lot about what goes on in his mind and in his emotional make up. Chi-squares and coefficients of correlation can ascertain the probabilities of how many people will prefer blue wall paper to red wall paper.

Then comes an actual situation: Joe Doaks can choose -- choose to work, or not to work, to cooperate, or to refuse to cooperate. Joe Doaks can even laugh when everyone else in the world cries. He can choose to be punished, when you would expect him to beg

to be let off. He can obey the law for 50 years, then murder someone. He can spend 37 years on the primrose path to damnation, and then shrink from crime, vice or from even the slightest indiscretion.

We are not implying that people are in no way predictable, nor do we impugn the psychological and the physical sciences that can throw considerable light on the how and why of human behavior. But people remain people, and they can choose "to be or not to be."

Suppose you said, "We know that the majority of people who end up in naval retraining commands come from broken homes, so let's exclude any recruit who comes from a broken home." But to do so might exclude some of our most successful personnel (because some of these have come from what one might call the worst possible type of home). Psychologists and psychiatrists are the last ones to claim that they can predict a man's future.

Is there any simple pencil and paper test or crucial question that will tell us whether someone is likely to go over the hill, or steal, or become a homosexual, once he has enlisted? We know of none to date, and efforts to develop one have been unsuccessful.

Even if a test seemed to work well enough to pay its cost, who will tell a "full-blooded American boy" that he is ineligible for naval service because his test scores suggest that he <u>might</u> some day become a drunkard, or that he might eventually fall into some other kind of delinquency?

Moreover, how big is our manpower supply? Could we "leave out" all the "potential troublemakers" even if it were possible to identify them more accurately? A Senate Subcommittee stated (October 1958) that one out of every five men who should be available for military service have records of arrest; a survey conducted by the Army and Air Force demonstrated that men with records of arrest generally account for far more courts-martial, cases of AWOL, and prison sentences than men without such records. Yet we have to have a certain minimum number of these men on-board even in "peacetime", while in all-out mobilization the number will skyrocket and selectivity will become much

less easy. As a further example, we find that the man most likely to end up in a retraining command is one who has failed to complete high school. Even so, the Navy and the other services must accept their "share" of men who have not completed high school.

The question of screening troublemakers is and will remain important because disciplinary problems cost millions of dollars and thousands of officer and man hours. Desirable as it is to have the best men, such an ideal cannot be reached. We have some of the finest people in the world -- and we have our problem makers. Justifiable though our complaint may be about the level of some recruits, once we have them, we have to do what we can for their good and for that of the Navy. Many of those who don't look promising can and have become the finest sailors in the fleet. We increase the chances of such success if we catch the potential troublemakers at the beginning of their enlistments.

SCREENING OUT POTENTIAL TROUBLEMAKERS 3

IF THE BEHAVIOR OF HUMAN-BEINGS WERE WHOLLY DETERMINED BY HEREDITY AND ENVIRONMENT, OR BY PURELY MECHANICAL OR AUTOMATIC LAWS, PREDICTION OF FUTURE SUCCESS OR FAILURE WOULD BE EASY. IF WE COULD SAY THAT HUMAN-BEINGS, WHEN STIMULATED IN A PARTICULAR FASHION, WOULD AUTOMATICALLY REACT 2 IN A GIVEN WAY, NOT ONLY THE PREDICTABILITY BUT THE CONTROL OF HUMAN BEHAVIOR WOULD LIKEWISE BE EASY. INDEED SOME SCHOOLS OF THOUGHT AND TEACHING HAVE GIVEN THIS IMPRESSION. ONE SUCH SCHOOL MIGHT LOOSELY BE TERMED "BEHAVIORISTIC" 4, ANOTHER "MECHANISTIC" 4. DEVELOPMENTS IN THE SCIENCE OF CYBERNETICS 3 (A COMPARATIVE STUDY OF COMPLEX MECHANICAL-ELECTRONIC COMMUNICATIONS-SYSTEMS 3, SUCH AS COMPUTERS, WITH THE CONTROL SYSTEM FORMED BY THE NERVOUS-SYSTEM AND THE BRAIN) HAVE LED SOME PEOPLE TO BELIEVE THAT THERE IS NO ESSENTIAL DIFFERENCE BETWEEN MEN AND HIGHLY REFINED "COMPUTERS" 3.

BUT THE FACT OF THE MATTER IS THAT NO MACHINE EVER INVENTED, PAST, PRESENT OR FUTURE, NO MATTER HOW COMPLEX ITS OPERATIONS, HAS HAD OR CAN HAVE A BUILT-IN PERSONAL-MORAL-RESPONSIBILITY 3. A MAN CAN KNOW ALL THE RULES. HE CAN STUDY BEHAVIORAL-PRINCIPLES 3 ALL HIS LIFE. IN TURN, HE CAN BE STUDIED IN EXHAUSTING DETAIL. HIS BLOOD CAN BE TESTED; X-RAYS 4 AND FLUOROSCOPES 4 CAN STRIP HIM OF EVERY PHYSIOLOGICAL SECRET. ELECTROENCEPHALOGRAMS 4, PSYCHOGALVANOMETERS 4, AND RORSCHACH 4 DIAGNOSES CAN TELL A LOT ABOUT WHAT GOES ON IN HIS MIND AND IN HIS EMOTIONAL MAKE-UP. CHI-SQUARES 4 AND COEFFICIENTS-OF-CORRELATION 4 CAN ASCERTAIN THE PROBABILITIES OF HOW MANY PEOPLE WILL PREFER BLUE WALL-PAPER TO RED WALL-PAPER.

THEN COMES AN ACTUAL SITUATION: JOE-DOAKS 1 CAN CHOOSE -- CHOOSE TO WORK, OR NOT TO WORK, TO COOPERATE, OR TO REFUSE TO COOPERATE. JOE-DOAKS CAN EVEN LAUGH WHEN EVERYONE ELSE IN THE WORLD CRIES. HE CAN CHOOSE TO BE PUNISHED, WHEN YOU WOULD EXPECT HIM TO BEG TO BE LET OFF. HE CAN OBEY THE LAW FOR 50 YEARS, THEN MURDER SOMEONE. HE CAN SPEND 37 YEARS ON THE PRIMROSE-PATH TO DAMNATION, AND THEN SHRINK FROM CRIME, VICE, OR EVEN THE SLIGHTEST INDISCRETION.

WE ARE NOT IMPLYING THAT PEOPLE ARE IN NO WAY PREDICTABLE, NOR DO WE IMPUGN THE PSYCHOLOGICAL AND THE PHYSICAL SCIENCES THAT CAN THROW CONSIDERABLE LIGHT ON THE HOW AND WHY OF HUMAN BEHAVIOR. BUT PEOPLE REMAIN PEOPLE, AND THEY CAN CHOOSE "TO BE OR NOT TO BE".

SUPPOSE YOU SAID, "WE KNOW THAT THE MAJORITY OF PEOPLE WHO END UP IN NAVAL-RETRAINING-COMMANDS 3 COME FROM BROKEN HOMES, SO LET'S EXCLUDE ANY RECRUIT WHO COMES FROM A BROKEN HOME." BUT TO DO SO MIGHT EXCLUDE SOME OF OUR 3 MOST SUCCESSFUL PERSONNEL (BECAUSE SOME OF THESE HAVE COME FROM WHAT ONE MIGHT CALL THE WORST POSSIBLE TYPE OF HOME). PSYCHOLOGISTS AND PSYCHIATRISTS ARE THE LAST ONES TO CLAIM THAT THEY CAN PREDICT A MAN'S FUTURE.

IS THERE ANY SIMPLE PENCIL-AND-PAPER TEST, OR CRUCIAL QUESTION THAT WILL TELL US 3 WHETHER SOMEONE IS LIKELY TO GO-OVER-THE-HILL 3, OR STEAL, OR BECOME A HOMOSEXUAL, ONCE HE HAS ENLISTED? WE 3 KNOW OF NONE TO DATE, AND EFFORTS TO DEVELOP ONE HAVE BEEN UNSUCCESSFUL.

EVEN IF A TEST SEEMED TO WORK WELL ENOUGH TO PAY ITS COST, WHO WILL TELL A "FULL-BLOODED AMERICAN BOY" THAT HE IS INELIGIBLE FOR NAVAL SERVICE BECAUSE HIS TEST SCORES SUGGEST THAT HE MIGHT SOME DAY BECOME A DRUNKARD, OR THAT HE MIGHT EVENTUALLY FALL INTO SOME OTHER KIND OF DELINQUENCY?

MOREOVER, HOW BIG IS OUR MANPOWER SUPPLY? COULD WE "LEAVE OUT" ALL THE "POTENTIAL TROUBLEMAKERS" EVEN IF IT WERE POSSIBLE TO IDENTIFY THEM MORE ACCURATELY? A SENATE SUBCOMMITTEE STATED (OCTOBER 1 1958) THAT ONE OUT OF EVERY FIVE MEN WHO SHOULD BE AVAILABLE FOR MILITARY SERVICE HAVE RECORDS OF ARREST; A SURVEY CONDUCTED BY THE ARMY AND AIR FORCE DEMONSTRATED THAT MEN WITH RECORDS OF ARREST GENERALLY ACCOUNT FOR FAR MORE COURTS-MARTIAL 3, CASES OF AWOL 3, AND PRISON SENTENCES THAN MEN WITHOUT SUCH RECORDS. YET WE HAVE TO HAVE A CERTAIN MINIMUM NUMBER OF THESE MEN ON-BOARD 3 EVEN IN "PEACETIME" WHILE IN ALL-OUT MOBILIZATION THE NUMBER WILL SKYROCKET, AND SELECTIVITY WILL BECOME MUCH LESS EASY. AS A FURTHER EXAMPLE, WE FIND THAT THE MAN MOST LIKELY TO END UP IN A RETRAINING-COMMAND 3 IS ONE WHO HAS FAILED TO COMPLETE HIGH-SCHOOL. EVEN SO, THE NAVY AND THE OTHER SERVICES MUST ACCEPT THEIR "SHARE" OF MEN WHO HAVE NOT COMPLETED HIGH-SCHOOL.

THE QUESTION OF SCREENING TROUBLEMAKERS IS AND WILL REMAIN IMPORTANT BECAUSE DISCIPLINARY PROBLEMS COST MILLIONS OF DOLLARS AND THOUSANDS OF OFFICER AND MAN HOURS. DESIRABLE AS IT IS TO HAVE THE BEST MEN, SUCH AN IDEAL CANNOT BE REACHED. WE HAVE SOME OF THE FINEST PEOPLE IN THE WORLD -- AND WE HAVE OUR PROBLEM-MAKERS 3. JUSTIFIABLE THOUGH OUR COMPLAINT MAY BE ABOUT THE LEVEL OF SOME RECRUITS, ONCE WE HAVE THEM, WE HAVE TO DO WHAT WE CAN FOR THEIR GOOD AND FOR THAT OF THE NAVY. MANY OF THOSE WHO DON'T LOOK PROMISING CAN AND HAVE BECOME THE FINEST SAILORS IN THE FLEET. WE INCREASE THE CHANCES OF SUCH SUCCESS IF WE CATCH THE POTENTIAL TROUBLEMAKERS AT THE BEGINNING OF THEIR ENLISTMENTS.

CLASS 3
AWOL 1
BEHAVIORAL-PRINCIPLES 1
COMMUNICATIONS-SYSTEMS 1
COMPUTERS 2
COURTS-MARTIAL 1
CYBERNETICS 1
GO-OVER-THE-HILL 1
NAVAL-RETRAINING-COMMANDS 1
ON-BOARD 1
OUR 4
PERSONAL-MORAL-RESPONSIBILITY 1
PROBLEM-MAKERS 1
RETRAINING-COMMAND 1
TROUBLEMAKERS 4
US 1
WE 15
NU 16 FQ 37

CLASS 4
BEHAVIORISTIC 1
CHI-SQUARES 1
COEFFICIENTS-OF-CORRELATION 1
ELECTROENCEPHALOGRAMS 1
FLUOROSCOPES 1
MECHANISTIC 1
PSYCHOGALVANOMETERS 1
RORSCHACH 1
X-RAYS 1
NU 9 FQ 9

CLASS 2
ABOUT 2
ACCEPT 1
ACCOUNT 1
ACCURATELY 1
ACTUAL 1
ALL-OUT 1
AMERICAN 1
ANOTHER 1
ANY 2
ARMY 1
ARREST 2
ASCERTAIN 1
AUTOMATIC 1
AUTOMATICALLY 1
AVAILABLE 1
BECAUSE 3
BECOME 4
BEGINNING 1
BEHAVIOR 3
BELIEVE 1
BETWEEN 1
BROKEN 2
BUILT-IN 1
CANNOT 1
CERTAIN 1
COMPARATIVE 1
COMPLAINT 1
COMPLETE 1
COMPLETED 1
COMPLEX 2
CONDUCTED 1
CONSIDERABLE 1
CONTROL 2
COOPERATE 2
CRUCIAL 1
DAMNATION 1
DELINQUENCY 1
DEMONSTRATED 1
DESIRABLE 1
DETAIL 1
DETERMINED 1
DEVELOP 1
DEVELOPMENTS 1
DIAGNOSES 1
DIFFERENCE 1
DISCIPLINARY 1
DOLLARS 1

DRUNKARD 1
EASY 3
EFFORTS 1
EMOTIONAL 1
ENLISTED 1
ENLISTMENTS 1
ENOUGH 1
ENVIRONMENT 1
ESSE[illegible]AL 1
EVEN 6
EVENTUALLY 1
EVER 1
EVERY 2
EVERYONE 1
EXAMPLE 1
EXCLUDE 2
EXHAUSTING 1
EXPECT 1
FAILURE 1
FASHION 1
FINEST 2
FULL-BLOODED 1
FURTHER 1
FUTURE 3
GENERALLY 1
GIVEN 2
HEREDITY 1
HIGH-SCHOOL 2
HIGHLY 1
HOMOSEXUAL 1
HUMAN 2
HUMAN-BEINGS 2
IDEAL 1
IDENTIFY 1
IMPLYING 1
IMPORTANT 1
IMPRESSION 1
IMPUGN 1
INCREASE 1
INDEED 1
INDISCRETION 1
INELIGIBLE 1
INTO 1
INVENTED 1
JUSTIFIABLE 1
LEVEL 1
LIKELY 2
LIKEWISE 1
LOOSELY 1
MACHINE 1
MAJORITY 1
MAKE-UP 1
MANPOWER 1
MANY 2
MATTER 2
MECHANICAL 1
MECHANICAL-ELECTRONIC 1
MILITARY 1
MILLIONS 1
MINIMUM 1
MOBILIZATION 1
MOREOVER 1
MURDER 1
NAVAL 1
NAVY 2
NERVOUS-SYSTEM 1
NUMBER 2
OBEY 1
OFFICER 1
ONLY 1
OPERATIONS 1
OTHER 2
PARTICULAR 1
PEACETIME 1
PENCIL-AND-PAPER 1
PEOPLE 7
PERSONNEL 1
PHYSICAL 1
PHYSIOLOGICAL 1
POSSIBLE 2
POTENTIAL 3
PREDICT 1
PREDICTABILITY 1
PREDICTABLE 1
PREDICTION 1
PREFER 1
PRESENT 1
PRIMROSE-PATH 1
PRISON 1
PROBABILITIES 1
PROBLEMS 1
PROMISING 1
PSYCHIATRISTS 1
PSYCHOLOGICAL 1
PSYCHOLOGISTS 1
PUNISHED 1
PURELY 1
QUESTION 2
REACT 1
RECORDS 3
RECRUIT 1
RECRUITS 1
REFINED 1
REFUSE 1
REMAIN 2
SAILORS 1
SECRET 1
SELECTIVITY 1
SENATE 1
SENTENCES 1
SERVICE 2
SERVICES 1
SIMPLE 1
SITUATION 1
SKYROCKET 1
SLIGHTEST 1
SOMEONE 2
STIMULATED 1
STUDIED 1
STUDY 2
SUBCOMMITTEE 1
SUCCESS 2
SUCCESSFUL 1
SUGGEST 1
SUPPLY 1
SUPPOSE 1
SURVEY 1
SYSTEM 1
THOUSANDS 1
UNSUCCESSFUL 1
WALL-PAPER 2
WHETHER 1
WHOLLY 1
WITHOUT 1
NU 181 FQ 235

CLASS 1

--	2
1958	1
37	1
50	1
A	17
AIR	1
ALL	3
AN	2
AND	24
ARE	3
AS	3
AT	1
BE	12
BEEN	1
BEG	1
BEST	1
BIG	1
BLOOD	1
BLUE	1
BOY	1
BRAIN	1
BUT	4
BY	4
CALL	1
CAN	18
CASES	1
CATCH	1
CHANCES	1
CHOOSE	4
CLAIM	1
COME	2
COMES	2
COST	2
COULD	2
CRIES	1
CRIME	1
DATE	1
DAY	1
DO	3
DON'T	1
ELSE	1
END	2
FACT	1
FAILED	1
FALL	1
FAR	1
FIND	1
FIVE	1
FLEET	1
FOR	6
FORCE	1
FORMED	1
FROM	4
GOES	1
GOOD	1
HAD	1
HAS	3
HAVE	15
HE	9
HIM	2
HIS	5
HOME	2
HOMES	1
HOURS	1
HOW	4
IF	5
IN	15
IS	9
IT	2
ITS	2
JOE-DOAKS	2
KIND	1
KNOW	3
LAST	1
LAUGH	1
LAW	1
LAWS	1
LEAVE	1
LED	1
LESS	1
LET	1
LET'S	1
LIFE	1
LIGHT	1
LOOK	1
LOT	1
MAN	3
MAN'S	1
MAY	1
MEN	7
MIGHT	5
MIND	1
MORE	2
MOST	2
MUCH	1
MUST	1
NO	4
NONE	1
NOR	1
NOT	5
OCTOBER	1
OF	31
OFF	1
ON	3
ONCE	2
ONE	5
ONES	1
OR	13
OUT	3
PAST	1
PAY	1
REACHED	1
RED	1
RULES	1
SAID	1
SAY	1
SCHOOL	1
SCHOOLS	1
SCIENCE	1
SCIENCES	1
SCORES	1
SCREENING	2
SEEMED	1
SHARE	1
SHOULD	1
SHRINK	1
SO	3
SOME	8
SPEND	1
STATED	1
STEAL	1
STRIP	1
SUCH	5
TEACHING	1
TELL	3
TERMED	1
TEST	3
TESTED	1
THAN	1
THAT	15
THE	38
THEIR	3
THEM	2
THEN	3
THERE	2
THESE	2
THEY	2
THIS	1
THOSE	1
THOUGH	1
THOUGHT	1
THROW	1
TO	26
TURN	1
TYPE	1
UP	2
VICE	1
WAY	2
WELL	1
WERE	2
WHAT	3
WHEN	3
WHILE	1
WHO	7
WHY	1
WILL	6
WITH	2
WORK	3
WORLD	2
WORST	1
WOULD	4
YEARS	2
YET	1
YOU	2

NU 174 FQ 538

7. Screening Out Potential Troublemakers --Revision

The problem of screening out potential troublemakers is and will remain important, because disciplinary problems cost the Navy millions of dollars and thousands of officer and man hours. It would be very desirable to develop and improve tests which would predict accurately whether a man seeking to enter the Navy would become a serious troublemaker, and thus enable the Navy to screen him out before entry.

But simple pencil and paper tests, or key questions, -- or even elaborate and expensive tests drawing deeply on the resources of physiology and psychology -- have not to date been successful in this sort of prediction. Accordingly, the Navy has not so far been able to tell ahead of time whether a particular man, after he has been accepted, will go over the hill, or steal, or become a drunkard or a homosexual, or be a troublemaker in other ways.

There seem to be two basic reasons for the lack of successful tests. First, the behavior of human beings even on the average is not fully determined by heredity, environment, or mechanical or automatic laws. This is true even though the psychological and physiological sciences throw considerable light on the how and why of human behavior.

Second, the individual human being is intensely variable. In an actual situation, Joe Doaks can choose to work, or not to work. He can laugh, when almost everyone else cries. He can choose to be punished, when you would expect him to beg to be let off. He can obey the law for 30 years, then murder someone.

Some predictive factors have been observed, such as a record of arrest, or coming from a broken home, or failure to complete high school. A survey conducted by the Army and Air Force demonstrated that men with records of arrest generally account for far more courts martial, absences without leave, and prison sentences than men without such records. In the Navy the men who are most likely to end up in naval retraining commands are those who have not completed high school.

But even so, the Navy cannot in practice be highly selective in its choice of men, for several reasons. First, some of our most successful men come from what might be called the worst possible type of home, or have had a record of arrest, or have failed to complete high school. To exclude such men would prevent such successes. Second, the manpower supply of the United States is not unlimited. The supply is not big enough for the Navy to screen out all potential troublemakers even if it were possible to identify them accurately. A Senate Subcommittee stated (October 1958) that one out of every five men who should be available for military service have records of arrest. Of course, in times of all-out mobilization the number of men that the Navy will have to accept will skyrocket.

So, justifiable as our complaint may be about the level of some of the Navy's recruits, once we have them, we have to do what we can for their good and the good of the Navy. Many of those who did not look promising in the past are now among the finest sailors in the fleet. We increase the chances of success like this if we treat disciplinary cases wisely when they occur at the beginning of enlistment.

SCREENING OUT POTENTIAL TROUBLEMAKERS

THE PROBLEM OF SCREENING OUT POTENTIAL TROUBLEMAKERS 3 IS AND WILL REMAIN IMPORTANT, BECAUSE DISCIPLINARY PROBLEMS COST THE NAVY MILLIONS OF DOLLARS AND THOUSANDS OF OFFICER AND MAN HOURS. IT WOULD BE VERY DESIRABLE TO DEVELOP AND IMPROVE TESTS WHICH WOULD PREDICT ACCURATELY WHETHER A MAN SEEKING TO ENTER THE NAVY WOULD BECOME A SERIOUS TROUBLEMAKER 3 , AND THUS ENABLE THE NAVY TO SCREEN HIM OUT BEFORE ENTRY.

BUT SIMPLE PENCIL-AND-PAPER TESTS, OR KEY QUESTIONS -- OR EVEN ELABORATE AND EXPENSIVE TESTS DRAWING DEEPLY ON THE RESOURCES OF PHYSIOLOGY AND PSYCHOLOGY -- HAVE NOT TO DATE BEEN SUCCESSFUL IN THIS SORT OF PREDICTION. ACCORDINGLY, THE NAVY HAS NOT SO FAR BEEN ABLE TO TELL AHEAD OF TIME WHETHER A PARTICULAR MAN, AFTER HE HAS BEEN ACCEPTED, WILL GO-OVER-THE-HILL 3, OR STEAL, OR BECOME A DRUNKARD OR A HOMOSEXUAL, OR BE A TROUBLEMAKER IN OTHER WAYS.

THERE SEEM TO BE TWO BASIC REASONS FOR THE LACK OF SUCCESFUL TESTS. FIRST, THE BEHAVIOR OF HUMAN-BEINGS, EVEN ON THE AVERAGE, IS NOT FULLY DETERMINED BY HEREDITY, ENVIRONMENT, OR MECHANICAL OR AUTOMATIC LAWS. THIS IS TRUE EVEN THOUGH THE PSYCHOLOGICAL AND PHYSIOLOGICAL SCIENCES THROW CONSIDERABLE LIGHT ON THE HOW AND WHY OF HUMAN BEHAVIOR. SECOND, THE INDIVIDUAL HUMAN-BEING IS INTENSELY VARIABLE. IN AN ACTUAL SITUATION JOE-DOAKS 1 CAN CHOOSE TO WORK, OR NOT TO WORK. HE CAN LAUGH, WHEN ALMOST EVERYONE ELSE CRIES. HE CAN CHOOSE TO BE PUNISHED, WHEN YOU WOULD EXPECT HIM TO BEG TO BE LET OFF. HE CAN OBEY THE LAW FOR 30 YEARS, THEN MURDER SOMEONE.

SOME PREDICTIVE FACTORS HAVE BEEN OBSERVED, SUCH AS A RECORD OF ARREST, OR COMING FROM A BROKEN HOME, OR FAILURE TO COMPLETE HIGH-SCHOOL. A SURVEY CONDUCTED BY THE ARMY AND THE AIR-FORCE DEMONSTRATED THAT MEN WITH RECORDS OF ARREST GENERALLY ACCOUNT FOR FAR MORE COURTS-MARTIAL 3, ABSENCES WITHOUT LEAVE, AND PRISON SENTENCES THAN MEN WITHOUT SUCH RECORDS. IN THE NAVY THE MEN WHO ARE MOST LIKELY TO END UP IN NAVAL-RETRAINING-COMMANDS 3 ARE THOSE WHO HAVE NOT COMPLETED HIGH-SCHOOL.

BUT EVEN SO THE NAVY CANNOT IN PRACTICE BE HIGHLY SELECTIVE IN ITS CHOICE OF MEN, FOR SEVERAL REASONS. FIRST, SOME OF OUR MOST SUCCESSFUL MEN COME FROM WHAT MIGHT BE CALLED THE WORST POSSIBLE TYPE OF HOME, OR HAVE HAD A RECORD OF ARREST, OR HAVE FAILED TO COMPLETE HIGH-SCHOOL. TO EXCLUDE SUCH MEN WOULD PREVENT SUCH SUCCESSES. SECOND, THE MANPOWER SUPPLY OF THE UNITED-STATES 1 IS NOT UNLIMITED. THE SUPPLY IS NOT BIG ENOUGH FOR THE NAVY TO SCREEN OUT ALL POTENTIAL TROUBLEMAKERS EVEN IF IT WERE POSSIBLE TO IDENTIFY THEM ACCURATELY. A SENATE SUBCOMMITTEE STATED (OCTOBER 1 1958) THAT ONE OUT OF EVERY FIVE MEN WHO SHOULD BE AVAILABLE FOR MILITARY SERVICE HAVE RECORDS OF ARREST. OF-COURSE, IN TIMES OF ALL-OUT MOBILIZATION THE NUMBER OF THE MEN THAT THE NAVY WILL HAVE TO ACCEPT WILL SKYROCKET.

SO, JUSTIFIABLE AS OUR COMPLAINT MAY BE ABOUT THE LEVEL OF SOME OF THE NAVY'S RECRUITS, ONCE WE HAVE THEM, WE HAVE TO DO WHAT WE CAN FOR THEIR GOOD AND FOR THE GOOD OF THE NAVY. MANY OF THOSE WHO DID NOT LOOK PROMISING IN THE PAST ARE NOW AMONG THE FINEST SAILORS IN THE FLEET. WE INCREASE THE CHANCES OF SUCCESS LIKE THIS IF WE TREAT DISCIPLINARY CASES WISELY WHEN THEY OCCUR AT THE BEGINNING OF ENLISTMENT.

CLASS 3
COURTS-MARTIAL 1
GO-OVER-THE-HILL 1
NAVAL-RETRAINING-COMMANDS 1
TROUBLEMAKER 2
TROUBLEMAKERS 3
NU 5 FQ 8

CLASS 2
ABLE 1
ABOUT 1
ABSENCES 1
ACCEPT 1
ACCEPTED 1
ACCORDINGLY 1
ACCOUNT 1
ACCURATELY 2
ACTUAL 1
AFTER 1
AHEAD 1
AIR-FORCE 1
ALL-OUT 1
ALMOST 1
AMONG 1
ARMY 1
ARREST 4
AUTOMATIC 1
AVAILABLE 1
AVERAGE 1
BASIC 1
BECAUSE 1
BECOME 2
BEFORE 1
BEGINNING 1
BEHAVIOR 2
BROKEN 1
CANNOT 1
COMPLAINT 1
COMPLETE 2
COMPLETED 1
CONDUCTED 1
CONSIDERABLE 1
DEEPLY 1
DEMONSTRATED 1
DESIRABLE 1
DETERMINED 1
DEVELOP 1
DISCIPLINARY 2
DOLLARS 1
DRUNKARD 1
ELABORATE 1
ENABLE 1
ENLISTMENT 1
ENOUGH 1
ENTER 1
ENTRY 1
ENVIRONMENT 1
EVEN 5
EVERY 1
EVERYONE 1
EXCLUDE 1
EXPECT 1
EXPENSIVE 1
FACTORS 1
FAILURE 1
FINEST 1
FULLY 1
GENERALLY 1
HEREDITY 1
HIGH-SCHOOL 3
HIGHLY 1
HOMOSEXUAL 1
HUMAN 1
HUMAN-BEING 1
HUMAN-BEINGS 1
IDENTIFY 1
IMPORTANT 1
IMPROVE 1
INCREASE 1
INDIVIDUAL 1
INTENSELY 1
JUSTIFIABLE 1
LEVEL 1
LIKELY 1
MANPOWER 1
MANY 1
MECHANICAL 1
MILITARY 1
MILLIONS 1
MOBILIZATION 1
MURDER 1
NAVY 9
NAVY'S 1
NUMBER 1
OBEY 1
OBSERVED 1
OCCUR 1
OF-COURSE 1
OFFICER 1
OTHER 1
PARTICULAR 1
PENCIL-AND-PAPER 1
PHYSIOLOGICAL 1
PHYSIOLOGY 1
POSSIBLE 2
POTENTIAL 3
PRACTICE 1
PREDICT 1
PREDICTION 1
PREDICTIVE 1
PREVENT 1
PRISON 1
PROBLEM 1
PROBLEMS 1
PROMISING 1
PSYCHOLOGICAL 1
PSYCHOLOGY 1
PUNISHED 1
QUESTIONS 1
REASONS 2
RECORD 2
RECORDS 3
RECRUITS 1
REMAIN 1
RESOURCES 1
SAILORS 1
SECOND 2
SELECTIVE 1
SENATE 1
SENTENCES 1
SERIOUS 1
SERVICE 1
SEVERAL 1
SIMPLE 1
SITUATION 1
SKYROCKET 1
SOMEONE 1
SUBCOMMITTEE 1
SUCCESFUL 1
SUCCESS 1
SUCCESSES 1
SUCCESSFUL 2
SUPPLY 2
SURVEY 1
THOUSANDS 1
UNLIMITED 1
VARIABLE 1
VERY 1
WHETHER 2
WISELY 1
WITHOUT 2
NU 142 FQ 176

CLASS 1

--	2	IF	2	THAN	1
1958	1	IN	10	THAT	3
30	1	IS	6	THE	35
A	11	IT	2	THEIR	1
ALL	1	ITS	1	THEM	2
AN	1	JOE-DOAKS	1	THEN	1
AND	12	KEY	1	THERE	1
ARE	3	LACK	1	THEY	1
AS	2	LAUGH	1	THIS	3
AT	1	LAW	1	THOSE	2
BE	9	LAWS	1	THOUGH	1
BEEN	4	LEAVE	1	THROW	1
BEG	1	LET	1	THUS	1
BIG	1	LIGHT	1	TIME	1
BUT	2	LIKE	1	TIMES	1
BY	2	LOOK	1	TO	19
CALLED	1	MAN	3	TREAT	1
CAN	5	MAY	1	TRUE	1
CASES	1	MEN	8	TWO	1
CHANCES	1	MIGHT	1	TYPE	1
CHOICE	1	MORE	1	UNITED-STATES	1
CHOOSE	2	MOST	2	UP	1
COME	1	NOT	8	WAYS	1
COMING	1	NOW	1	WE	5
COST	1	OCTOBER	1	WERE	1
CRIES	1	OF	26	WHAT	2
DATE	1	OFF	1	WHEN	3
DID	1	ON	3	WHICH	1
DO	1	ONCE	1	WHO	4
DRAWING	1	ONE	1	WHY	1
ELSE	1	OR	13	WILL	4
END	1	OUR	2	WITH	1
FAILED	1	OUT	5	WORK	2
FAR	2	PAST	1	WORST	1
FIRST	2	SCIENCES	1	WOULD	5
FIVE	1	SCREEN	2	YEARS	1
FLEET	1	SCREENING	2	YOU	1
FOR	8	SEEKING	1		
FROM	2	SEEM	1		
GOOD	2	SHOULD	1		
HAD	1	SO	3		
HAS	2	SOME	3		
HAVE	9	SORT	1		
HE	4	STATED	1		
HIM	2	STEAL	1		
HOME	2	SUCH	4		
HOURS	1	TELL	1		
HOW	1	TESTS	4		

NU 133 FQ 366

8. Communication is Subtle

The following is taken from "The Talk of the Town" in "The New Yorker" for March 3, 1951.

A publisher in Chicago has sent us a pocket calculating machine by which we may test our writing to see whether it is intelligible. The calculator was developed by General Motors, who, not satisfied with giving the world a Cadillac, now dream of bringing perfect understanding to men. The machine (it is simply a celluloid card with a dial) is called the Reading-Ease Calculator and shows four grades of "reading ease" — Very Easy, Easy, Hard, and Very Hard. You count your words and syllables, set the dial, and an indicator lets you know whether anybody is going to understand what you have written. An instruction book came with it, and after mastering the simple rules we lost no time in running a test on the instruction book itself, to see how <u>that</u> writer was doing. The poor fellow! His leading essay, the one on the front cover, tested Very Hard.

Our next step was to study the first phrase on the face of the calculator: "How to test Reading-Ease of written matter." There is, of course, no such thing as reading ease of written matter. There is the ease with which matter can be read, but that is a condition of the reader, not of the matter. Thus the inventors and distributors of this calculator get off to a poor start, with a Very Hard instruction book and a slovenly phrase. Already they have one foot caught in the brier patch of English usage.

Not only did the author of the instruction book score badly on the front cover, but inside the book he used the word "personalize" in an essay on how to improve one's writing. A man who likes the word "personalize" is entitled to his choice, but we wonder whether he should be in the business of giving advice to writers. "Whenever possible," he wrote, "personalize your writing by directing it to the reader." As for us, we would as lief Simonize our grandmother as personalize our writing.

In the same envelope with the calculator, we received another training aid for writers -- a booklet called "How to Write Better," by Rudolf Flesch. This, too, we studied, and it quickly demonstrated the broncolike ability of the English language to throw whoever leaps cocksurely into the saddle. The language not only can toss a rider but knows a thousand tricks for tossing him, each more gay than the last. Dr. Flesch stayed in the saddle only a moment or two. Under the heading "Think Before You Write," he wrote, "The main thing to consider is your purpose in writing. Why are you sitting down to write?" And echo answered: Because, sir, it is more comfortable than standing up.

Communication by the written word is a subtler (and more beautiful) thing than Dr. Flesch or General Motors imagines. They contend that the "average reader" is capable of reading only what tests Easy, and that the writer should write at or below this level. This is a presumptuous and degrading idea. There is no average reader, and to reach down toward this mythical character is to deny that each of us is on the way up, is ascending. ("Ascending," by the way, is a word Dr. Flesch advises writers to stay away from. Too unusual.)

It is our belief that no writer can improve his work until he discards the dulcet notion that the reader is feeble-minded, for writing is an act of faith, not a trick of grammar. Ascent is at the heart of the matter. A country whose writers are following a calculating machine downstairs is not ascending -- if you will pardon the expression -- and a writer who questions the capacity of the person at the other end of the line is not a writer at all, merely a schemer. The movies long ago decided that a wider communication could be achieved by a deliberate descent to a lower level, and they walked proudly down until they reached the cellar. Now they are groping for the light switch, hoping to find the way out.

We have studied Dr. Flesch's instructions diligently, but we return for guidance in these matters to an earlier American, who wrote with more patience, more confidence,

"I fear chiefly," he wrote, "lest my expression may not be extra-vagant enough, may not wander far enough beyond the narrow limits of my daily experience, so as to be adequate to the truth of which I have been convinced.....Why level downward to our dullest perception always, and praise that as common sense? The commonest sense is the sense of men asleep, which they express by snoring."

Run that through your calculator! It may come out Hard, it may come out Easy. But it will come out whole, and it will last forever.

A PUBLISHER IN CHICAGO 1 HAS SENT US A POCKET CALCULATING MACHINE BY WHICH WE MAY TEST OUR WRITING TO SEE WHETHER IT IS INTELLIGIBLE. THE CALCULATOR WAS DEVELOPED BY GENERAL-MOTORS 1, WHO, NOT SATISFIED WITH GIVING THE WORLD A CADILLAC 4, NOW DREAM OF BRINGING PERFECT UNDERSTANDING TO MEN. THE MACHINE (IT IS SIMPLY A CELLULOID CARD WITH A DIAL) IS CALLED THE READING-EASE-CALCULATOR 3 AND SHOWS FOUR GRADES OF "READING-EASE" 3 -- VERY-EASY 3, EASY 3, HARD 3, AND VERY-HARD 3. YOU COUNT YOUR WORDS AND SYLLABLES, SET THE DIAL, AND AN INDICATOR LETS YOU KNOW WHETHER ANYBODY IS GOING TO UNDERSTAND WHAT YOU HAVE WRITTEN. AN INSTRUCTION BOOK CAME WITH IT, AND AFTER MASTERING THE SIMPLE RULES WE LOST NO TIME IN RUNNING A TEST ON THE INSTRUCTION BOOK ITSELF, TO SEE HOW THAT WRITER WAS DOING. THE POOR FELLOW! HIS LEADING ESSAY, THE ONE ON THE FRONT COVER, TESTED VERY-HARD.

NOT ONLY DID THE AUTHOR OF THE INSTRUCTION BOOK SCORE BADLY ON THE FRONT COVER, BUT INSIDE THE BOOK HE USED THE WORD "PERSONALIZE" IN AN ESSAY ON HOW TO IMPROVE ONE'S 1 WRITING. A MAN WHO LIKES THE WORD "PERSONALIZE" IS ENTITLED TO HIS CHOICE, BUT WE WONDER IF HE SHOULD BE IN THE BUSINESS OF GIVING ADVICE TO WRITERS. "WHENEVER POSSIBLE," HE WROTE, "PERSONALIZE YOUR WRITING BY DIRECTING IT TO THE READER." AS FOR US, WE WOULD AS LIEF 2 SIMONIZE OUR GRANDMOTHER AS PERSONALIZE OUR WRITING.

IN THE SAME ENVELOPE WITH THE CALCULATOR, WE RECEIVED ANOTHER TRAINING-AID FOR WRITERS -- A BOOKLET CALLED "HOW TO WRITE BETTER", BY RUDOLF-FLESCH 1. THIS TOO, WE STUDIED, AND IT QUICKLY DEMONSTRATED THE BRONCOLIKE ABILITY OF THE ENGLISH LANGUAGE TO THROW WHOEVER LEAPS COCKSURELY INTO THE SADDLE. THE LANGUAGE NOT ONLY CAN TOSS A RIDER BUT KNOWS A THOUSAND TRICKS FOR TOSSING HIM, EACH MORE GAY THAN THE LAST. DR-FLESCH 1 STAYED IN THE SADDLE ONLY A MOMENT OR TWO. UNDER THE HEADING "THINK BEFORE YOU WRITE," HE WROTE, "THE MAIN THING TO CONSIDER IS YOUR PURPOSE IN WRITING. WHY ARE YOU SITTING DOWN TO WRITE?" AND ECHO ANSWERED: BECAUSE, SIR, IT IS MORE COMFORTABLE THAN STANDING UP.

COMMUNICATION BY THE WRITTEN WORD IS A SUBTLER (AND MORE BEAUTIFUL THING THAN DR-FLESCH OR GENERAL-MOTORS IMAGINES. THEY CONTEND THAT THE "AVERAGE-READER" 3 IS CAPABLE OF READING ONLY WHAT TESTS EASY, AND THAT THE WRITER SHOULD WRITE AT OR BELOW THIS LEVEL. THIS IS A PRESUMPTUOUS AND DEGRADING IDEA. THERE IS NO AVERAGE-READER, AND TO REACH DOWN TOWARD THIS MYTHICAL CHARACTER IS TO DENY THAT EACH OF US IS ON THE WAY UP, IS ASCENDING. ("ASCENDING", BY THE WAY, IS A WORD DR-FLESCH ADVISES WRITERS TO STAY AWAY FROM. TOO UNUSUAL.)

IT IS OUR BELIEF THAT NO WRITER CAN IMPROVE HIS WORK UNTIL HE DISCARDS THE DULCET NOTION THAT THE READER IS FEEBLE-MINDED, FOR WRITING IS AN ACT OF FAITH, NOT A TRICK OF GRAMMAR. ASCENT IS AT THE HEART OF THE MATTER. A COUNTRY WHOSE WRITERS ARE FOLLOWING A CALCULATING MACHINE DOWNSTAIRS IS NOT ASCENDING -- IF YOU WILL PARDON THE EXPRESSION -- AND A WRITER WHO QUESTIONS THE CAPACITY OF THE PERSON AT THE OTHER END OF THE LINE IS NOT A WRITER AT ALL, MERELY A SCHEMER. THE MOVIES LONG AGO DECIDED THAT A WIDER COMMUNICATION COULD BE ACHIEVED BY A DELIBERATE DESCENT TO A LOWER LEVEL, AND THEY WALKED PROUDLY DOWN UNTIL THEY REACHED THE CELLAR. NOW THEY ARE GROPING FOR THE LIGHT SWITCH, HOPING TO FIND THE WAY OUT.

WE HAVE STUDIED DR-FLESCH'S 1 INSTRUCTIONS DILIGENTLY, BUT WE RETURN FOR GUIDANCE TO AN EARLIER AMERICAN 1, WHO WROTE WITH MORE PATIENCE, MORE CONFIDENCE. "I FEAR CHIEFLY," HE WROTE, "LEST MY EXPRESSION MAY NOT BE EXTRA-VAGANT 3 ENOUGH, MAY NOT WANDER FAR ENOUGH BEYOND 2 THE NARROW LIMITS OF MY DAILY EXPERIENCE, SO AS TO BE ADEQUATE TO THE TRUTH OF WHICH I HAVE BEEN CONVINCED .... WHY LEVEL DOWNWARD TO OUR DULLEST PERCEPTION ALWAYS, AND PRAISE THAT AS COMMON SENSE? THE COMMONEST SENSE IS THE SENSE OF MEN ASLEEP, WHICH THEY EXPRESS BY SNORING."

RUN THAT THROUGH YOUR CALC[illegible]FOR! IT MAY COME OUT HARD, IT MAY COME OUT EASY. BUT IT WILL COME OUT WHOLE, AND IT WILL LAST FOREVER.

CLASS 3
AVERAGE-READER 2
EASY 3
EXTRA-VAGANT 1
HARD 2
READING-EASE 1
READING-EASE-CALCULATOR 1
VERY-EASY 1
VERY-HARD 2
NO 8 FO 13

CLASS 2
ABILITY 1
ACHIEVED 1
ADEQUATE 1
ADVICE 1
ADVISES 1
AFTER 1
AGO 1
ALWAYS 1
ANOTHER 1
ANSWERED 1
ANYBODY 1
ASCENDING 3
ASCENT 1
ASLEEP 1
AUTHOR 1
AWAY 1
BADLY 1
BEAUTIFUL 1
BECAUSE 1
BEFORE 1
BELIEF 1
BELOW 1
BETTER 1
BEYOND 1
BOOKLET 1
BRONCOLIKE 1
BUSINESS 1
CALCULATING 2
CALCULATOR 3
CAPABLE 1
CAPACITY 1
CELLAR 1
CELLULOID 1
CHARACTER 1
CHIEFLY 1
COCKSURELY 1
COMFORTABLE 1
COMMON 1
COMMONEST 1
COMMUNICATION 2
CONFIDENCE 1
CONSIDER 1
CONTEND 1
CONVINCED 1
COUNTRY 1
COVER 2
DAILY 1
DECIDED 1
DEGRADING 1
DELIBERATE 1
DEMONSTRATED 1
DENY 1
DESCENT 1
DEVELOPED 1
DILIGENTLY 1
DIRECTING 1
DISCARDS 1
DOWNSTAIRS 1
DOWNWARD 1
DULCET 1
DULLEST 1
EARLIER 1
ECHO 1
ENGLISH 1
ENOUGH 2
ENTITLED 1
ENVELOPE 1
ESSAY 2
EXPERIENCE 1
EXPRESS 1
EXPRESSION 2
FEEBLE-MINDED 1
FELLOW 1
FOLLOWING 1
FOREVER 1
GRAMMAR 1
GRANDMOTHER 1
GUIDANCE 1
IDEA 1
IMAGINES 1
IMPROVE 2
INDICATOR 1
INSIDE 1
INSTRUCTION 3
INSTRUCTIONS 1
INTELLIGIBLE 1
INTO 1
ITSELF 1
LANGUAGE 2
LEVEL 3
LIEF 1
LIMITS 1
LOWER 1
MACHINE 3
MASTERING 1
MATTER 1
MERELY 1
MOMENT 1
MOVIES 1
MYTHICAL 1
NARROW 1
NOTION 1
ONE'S 1
ONLY 4
OTHER 1
PARDON 1
PATIENCE 1
PERCEPTION 1
PERFECT 1
PERSON 1
PERSONALIZE 4
POCKET 1
POSSIBLE 1
PRESUMPTUOUS 1
PROUDLY 1
PUBLISHER 1
PURPOSE 1
QUESTIONS 1
QUICKLY 1
READER 2
RECEIVED 1
RETURN 1
RIDER 1
SADDLE 2
SATISFIED 1
SCHEMER 1
SIMONIZE 1
SIMPLE 1
SIMPLY 1
STUDIED 2
SUBTLER 1
SYLLABLES 1
THOUSAND 1
TOWARD 1
TRAINING-AID 1
UNDER 1
UNDERSTAND 1
UNDERSTANDING 1
UNTIL 2
UNUSUAL 1
WANDER 1
WHENEVER 1
WHETHER 2
WHOEVER 1
WIDER 1
WONDER 1
WRITER 5
WRITERS 4
WRITTEN 2
NO 149 FO 186

CLASS	1	HIS	3	SO	1
--	4	HOPING	1	STANDING	1
A	23	HOW	3	STAY	1
ACT	1	I	2	STAYED	1
ALL	1	IF	2	SWITCH	1
AMERICAN	1	IN	7	TEST	2
AN	5	IS	22	TESTED	1
AND	15	IT	11	TESTS	1
ARE	3	KNOW	1	THAN	3
AS	5	KNOWS	1	THAT	9
AT	4	LAST	2	THE	50
BE	4	LEADING	1	THERE	1
BEEN	1	LEAPS	1	THEY	5
BOOK	4	LEST	1	THING	2
BRINGING	1	LETS	1	THINK	1
BUT	5	LIGHT	1	THIS	4
BY	8	LIKES	1	THROUGH	1
CALLED	2	LINE	1	THROW	1
CAME	1	LONG	1	TIME	1
CAN	2	LOST	1	TO	21
CARD	1	MAIN	1	TOO	2
CHICAGO	1	MAN	1	TOSS	1
CHOICE	1	MAY	5	TOSSING	1
COME	3	MEN	2	TRICK	1
COULD	1	MORE	5	TRICKS	1
COUNT	1	MY	2	TRUTH	1
DIAL	2	NO	3	TWO	1
DID	1	NOT	8	UP	2
DOING	1	NOW	2	US	3
DOWN	3	OF	15	USED	1
DR-FLESCH	3	ON	5	WALKED	1
DR-FLESCH'S	1	ONE	1	WAS	2
DREAM	1	OR	3	WAY	3
EACH	2	OUR	5	WE	8
END	1	OUT	4	WHAT	2
FAITH	1	POOR	1	WHICH	3
FAR	1	PRAISE	1	WHO	4
FEAR	1	REACH	1	WHOLE	1
FIND	1	REACHED	1	WHOSE	1
FOR	6	READING	1	WHY	2
FOUR	1	RUDOLF-FLESCH	1	WILL	3
FROM	1	RULES	1	WITH	5
FRONT	2	RUN	1	WORD	4
GAY	1	RUNNING	1	WORDS	1
GENERAL-MOTORS	2	SAME	1	WORK	1
GIVING	2	SCORE	1	WORLD	1
GOING	1	SEE	2	WOULD	1
GRADES	1	SENSE	3	WRITE	4
GROPING	1	SENT	1	WRITING	6
HAS	1	SET	1	WROTE	4
HAVE	3	SHOULD	2	YOU	6
HE	6	SHOWS	1	YOUR	4
HEADING	1	SIR	1		
HEART	1	SITTING	1		
HIM	1	SNORING			

NU 161 FQ 487

Appendix 3

"Giving Power to Words" by Philip W. Swain

The following article is reprinted with permission from the "American Journal of Physics", Vol. 13, No. 5, pages 318-320, October 1945.

GIVING POWER TO WORDS

by Philip W. Swain

Editor of Power, New York 18, N.Y.

"Fools rush in where angels fear to tread." Perhaps that is why I am not averse to giving physicists advice on how to write and edit. Within the past year I have told school mathematics teachers how they should teach, engineering deans what they should do about English studies, mechanical engineering professors what subjects they should teach and how.

Now I am at it again. Sooner or later this sort of thing may lead to my public undoing by some real expert.

So let me disarm you with a bit of frankness. I am not an expert writer. I feel like one of my engineering students at Yale many years ago. He wrote his "no help" pledge this way on his examination paper: "Mr. Swain, I give you my word of honor that I have neither given nor received help in this examination, of which the foregoing pages are ample evidence."

Aside from an invitation to do so, my only excuse for writing this paper is that I have been struggling with the practical problems of technical writing and editing for 23 years, and came to the editor's job with a rather broad background that

included a lot of Greek and Latin study, an academic major in physics, tutoring in physics, industrial factory jobs as machinist and steam fitter, two degrees in mechanical engineering, teaching and practicing mechanical engineering, serving as an army artillery officer and as a salesman.

Today I am convinced of what I long ago suspected -- that both physics and English are superlatively useful studies. Physics, the most fundamental of all engineering fundamentals, was my first love, and is still my favorite study.

Your work as teachers of physics in an age of applied science is of first importance. To do your job right you will need skill with three languages: the language of words, the language of mathematics and the language of pictures. Here we talk about the words.

KNOW YOUR READER AND YOUR OBJECTIVE

The first rule of good writing and good editing is to know your reader -- who he is, how he thinks; what words he uses, what words he understands; what interests him, what bores him; wherein he is smart, wherein stupid.

The second rule is to know what you wish to accomplish in the way of giving pleasure, causing amusement, winning friendship, imparting knowledge, awakening interests and reforming attitudes.

You are physicists and teachers of physics, so I assume that you write mostly for other physicists, for scientists in related fields, for students of elementary physics, for students of advanced physics and for the nonscientific public. More often than not, I assume, your aim will be to impart some understanding of physical laws, facts, methods or points of view, together with some share of your own enthusiasm.

So you physicist writers start by knowing just what you want to do, and to whom. There remains only "How?", the sixty-four dollar question. It is a question I cannot answer, nor anybody else. I can only pass along a few tips, and urge you to build skill by constant writing, testing and rewriting.

It will not be easy. English, the most gloriously illogical major language, is

better for poets than for physicists. But you will always find the right word if you look long enough.

MANY LANGUAGES

Note the many languages within our language. The college freshman learns that "the moment of a force about any specified axis is the product of the force and the perpendicular distance from the axis to the line of action of the force." Viewing the same physical principle, the engineer says: "To lift a heavy weight with a lever, a man should apply his strength to the end of a long lever arm and work the weight on a short lever arm." Out on the factory floor the foreman shouts, "Shove that brick up snug under the crowbar and get a good purchase; the crate is heavy." The salesman says: "Why let your men kill themselves heaving those boxes all day long? The job's easy with this new long-handled pinch bar. With today's high wages you'll save the cost the first afternoon."

Who can say which of these is the best English, or the worst? Each seems well suited to its purpose and audience, therefore good English, according to my lights.

May I venture to suggest a useful writing exercise for grown-up physicists. First write a sedate and technical little treatise on the gyroscope -- say, 1000 words. You might entitle it "A concise summary of the physical principles underlying gyroscopic phenomena." Then rewrite (and retitle) the piece seven times for the following seven types of audience or reader:

The Latin faculty of your university.

Some imaginary university president to be impressed by your profundity.

A student of first-year physics.

The Reader's Digest.

A Rotary Club meeting.

A mechanic, skilled but unschooled.

An eager 10-year-old boy, interested in gadgets.

TEN "RULES"

Now, at the risk of appearing utterly inconsistent, I am going to set down 10 "rules" of good writing. It will be fair to ask me how there can be general rules if it is true that each type of reader requires special treatment. Somewhat lamely, I answer that certain ways of writing (as distinguished from content) seem to go over well with the majority of American readers of all classes. If you follow these 10 suggestions blindly you will not always be right, but you will be right more often than you are wrong:

1. Short words are better than long.
2. Short sentences are better than long.
3. Short paragraphs are better than long.
4. Short articles are better than long.
5. Direct statement is better than indirect.
6. The active mood is better than the passive.
7. Don't pussyfoot.
8. Be simple, human and concise -- not complex, pompous and verbose.
9. Don't overwork "is", "was" and other parts of "to be".
10. To gain power, chop off Latin roots wherever Anglo-Saxon words can tell the same story. Thus, "They considered it improbable that circumstances would permit him to divulge the occurrence," would better be, "They didn't think he had a chance to tell the news."

Thus baldly set down without a lot of justifying "ifs" and "buts," these so-called rules may expose me to a barrage of exceptions from literary and scientific people. I grant the exceptions, but still insist that those who can write by those "rules" when they want to can best be trusted to break them wisely.

I must not appear to criticize the writings of one trained physicist to another, or the language these two understand best even if some words are both long and Latin.[1]

---

[1]For an analysis of such writings, see D. Koller, Am.J. Physics 13,99 (1945).

Two men of one mind talking shop are a pretty sight, whether they be physicists, machinists or baseball fans. They understand each other, and that is all that counts.

But let me ask you how many physicists can go through life with no fair opportunity to instruct or inspire some of the 99.997 percent of Americans who are not physicists. If we believe in the power and glory of physics, can we justify wasting such opportunities or use the lame excuse that the other fellow is stupid?

Learning to talk and write so the other fellow can understand you is a prime duty of every man with any wisdom worth sharing. The rules I have listed apply with particular force wherever the subject discussed is strange to the reader.

Whether you agree or not, I urge you to try this experiment: Hunt up some bit of expository writing full of long words, Latin roots, long sentences, long paragraphs and the passive mood. Rewrite by the "rules", and test the result. Is it easier to read? Does it save the reader's time? Does the writing have more power? What scientific or literary values, if any, have been lost?

SHAKESPEARE REWRITTEN

Out of curiosity, I reversed this procedure to see if it would make good writing bad. It did. My starting point was part of Mark Antony's speech to the mob in Julius Caesar. Rome might seem a good place to use Latin words, but Shakespeare knew better and made Mark speak Anglo-Saxon because there was a job to be done and no time for fooling around. Note also that Mark Antony shows no fear of direct statements, personal pronouns or one-syllable words. I now quote as Shakespeare wrote:

> Good friends, sweet friends, let me not stir you up
> to such a sudden flood of mutiny.
> They that have done this deed are honourable:
> What private griefs they have, alas! I know not,
> What made them do it; they are wise and honourable,
> And will, no doubt, with reasons answer you.
> I come not, friends, to steal away your hearts:

I am no orator, as Brutus is;
But, as you know me all, a plain blunt man,
That love my friends; and that they know full well
That gave me public leave to speak of him.
For I have neither wit, nor words, nor worth,
Action, nor utterance, nor the power of speech,
To stir men's blood; I only speak right on . . . .

Now I will rewrite into this speech the faults of the typical engineering manuscript -- long words, long sentences, pussyfooting expressions, passive moods, failure to call the other fellow "you" and yourself "I". Note the difference!

It is not the intention of the speaker to create in the minds of the friends and other gentlemen present any rapid increase in antagonistic and violent emotions. The persons who sustain the responsibility for this action are gentlemen of substantial reputation. It has not been feasible for the speaker to determine what personal grievances may have impelled them to concur in the action under discussion. However, due to the fact that they are intelligent and of satisfactory reputation, it may be assumed that they will stand prepared to present apparently defensible explanations of their procedure.

It should not be considered to be the intention of the speaker in appearing before you to influence your emotions in such a way as to advance his own personal selfish interests. The speaker is not properly what might be termed an adept in the profession of public speaking, as might be properly stated of Mr. Brutus. It is perhaps not unreasonable to make the assumption that all of the gentlemen here present are acquainted with the fact that the speaker is a person of uncomplicated character and one not addicted to circumlocutions and other types of round-about operations -- also that an understandable fondness was maintained by the speaker toward this rather close acquaintance.

Circumstances such as these are entirely familiar to those gentlemen who have accorded the speaker permission to present comments regarding this person. As far as the personal qualifications of the speaker are concerned, his abilities do not include the intelligence, the vocabulary or the character, the procedure, the verbal delivery or the skilfulness in enunciation requisite to the creation of excessive excitement in an audience. To speak with entire accuracy, it is practically impossible for the speaker to accomplish more than an unpretentious enumeration of the circumstances of a situation in correct sequence.

ENGLISH VERSUS JARGON

Here is a typical engineer's sentence of 39 words, most of which are wasted.

> It is assumed ordinarily that in computations of this character it is desirable to arrange the various elements of the problem in the form of a tabulation in order to insure that avoidable errors are reduced to a minimum.

Why not say:

> Most engineers like to tabulate such data to reduce errors.

Lampooning the pompous jargon of doctors, lawyers, engineers and other professionals, C. M. Ripley, of the General Electric Company, quotes an economist:

> The significance of this widespread problem can be grasped when it is recognized that the phenomenal growth of cities which has taken place during the past several decades is attributable to our efforts, in this industrial era, to achieve convenient physical accessibility between the many interdependent parts of our intricate economic and social mechanism.

The economist is trying to tell us that cities have grown because of good transit systems.

Ripley shows how short words bring power to the King James version of the Bible:

> Rise up, my love, my fair one, and come away. For lo, the winter is past, the rain is over, and gone.

From all English literature before Shakespeare, Ripley reminds us, Bartlett lists 363 familiar quotations averaging 11 words each -- and three words out of four have but one syllable.

CONCLUSION

In closing let me sum up in this way: (1) know your reader; (2) know your objective; (3) be simple, direct and concise.

Appendix 4

Bibliography

The following short bibliography is divided into four parts. The sequence of entries in each part is roughly from more simple to less simple. I hope these references will be of use to a reader who wishes to pursue the subject of explanation.

1. Explanation, Exposition, and English Composition

Flesch, Rudolf / The Art of Plain Talk / Harper and Bros., New York, N.Y. / 1946, 210 pp

This famous book caused a revolution in the writing of ordinary English, and has sold well over 100,000 copies. It contains a simple readability scale, a way of measuring the ease or difficulty of reading something written. A great many writers have applied this scale, and have obtained useful results. The book is easy to read, interesting, entertaining, and worthwhile.

Borden, Richard C. / Public Speaking -- As Listeners Like It / Harper and Bros., New York, N.Y. / 1935, 111 pp

This is an excellent guide for those who wish to speak well -- and for those who wish to explain well. It has four chapters; "Listeners' Laws" for "Speech Organization", "Speech Substance", "Speech Phraseology," and "Speech Delivery", respectively.

Flesch, Rudolf / The Art of Readable Writing / Harper and Bros., New York, N.Y. / 1949, 237 pp

The purpose of this book is "to help you in your writing", to put to use in

people's writing the "enormous amount of recent scientific information" about English grammar, usage and composition, to provide a "modern scientific rhetoric that you can apply to your own writing." The 21 chapters include: You and Aristotle, The Importance of Being Trivial, Degrees of Plain Talk, How to Operate a Blue Pencil, etc. This book is full of insights, pointed yet amusing remarks, and useful information.

Strunk, William, Jr., and E. B. White / The Elements of Style / The Macmillan Co., New York, N.Y. / 1962, paperback, 71 pp

This was first published in 1919 by William Strunk, Jr., a professor of English at Cornell University, Ithaca, N.Y.; it was republished in 1959 "with revisions, an introduction, and a new chapter on writing" by E. B. White, a contributor to the New Yorker, and a pupil of Professor Strunk's.

Wm. Strunk: "This book aims to give in brief space the principal requirements of plain English style."

E. B. White: "I shall have a word or two to say about attitudes in writing: the how, the why, the beartraps, the power, and the glory."

A valuable little book.

Gunning, Robert / The Technique of Clear Writing / McGraw Hill Book Co., New York, N.Y. / 1952, 289 pp

The author has had eight years experience as a business consultant in improving readability. Part 1 includes as chapters: The Fight Against Fog, What Every Writer Should Know about Readers, Readability Yardsticks. Part 2 explains ten principles, from no. 1, "Keep sentences short", to no. 10 "Write to express not to impress".

Berkeley, Edmund C. / How to Explain Clearly / Berkeley Enterprises Inc., 815 Washington St., Newtonville, Mass., 02160 / 1959, 23 pp

A discussion and listing of over 40 factors affecting explanation. This was the source for most of Chapter 4 in this book.

Perrin, Porter G. / Writer's Guide and Index to English / Scott Foresman and Co., Chicago, Ill. / 1950, hardbound, 833 pp

This book "describes our current language in a pattern planned for use in college courses in composition, and for "encouraging students in effective communication of facts and ideas". The chapters include: The Activity of Writing, Varieties of English, Writing Paragraphs, the Meaning of Words, etc. The book is full of indispensable information.

Rorabacher, Louise E. / Assignments in Exposition / Harper and Bros., New York, N.Y. / 1946 / 374 pp

A useful book, designed for teaching students in English composition courses; it is written clearly and directly:

> The emphasis here is not on creative writing as an art but on practical composition as a means of communicating ideas quickly and clearly in whatever form the workaday world may require .... that useful kind of expression most frequentiy demanded of the average man or woman which is known as infor ative or expository writing.

Tanner, William N. / Composition and Rhetoric / Ginn and Co., Boston, Mass . / 1922, 539 pp

Chapter 17, "Exposition", begins:

> Whenever we define a word, tell our companions how to play a game, recommend a book to a friend, ... or explain a law of science, we employ exposition. We may define exposition, then, as that form of composition by means of which we give explanations of facts, ideas, methods, and principles. The purpose of exposition is to make our meaning clear to the understanding of our hearers and readers. It appeals chiefly to the mind, rarely to the feelings. Next to narration, exposition is the most common form of discourse. The practice of

exposition develops observation, judgment, orderly thinking, and accurate expression. It is of the utmost importance that we learn to explain to others exactly what we mean.

Although this book is in many ways old-fashioned, and most of the chapters deal only indirectly with exposition, the book contains much useful and pithy information for improving explanation.

Berkeley, Edmund C. / "Supplement 1: Words and Ideas" in "Giant Brains or Machines that Think" / John Wiley and Sons, New York, N.Y. / 1949, 270 pp

A discussion of words, ideas, and explanation. In this book four classes of words for explaining were distinguished, and identified as Classes 1, 2, 3, and 4, and they were used and controlled to produce better explanation.

2. Technical Writing in Business and Engineering

Schutte, William M., and Erwin R. Steinberg / Communication in Business and Industry / Hold, Rinehart, and Winston, New York / 1965, hardbound, 393 pp, $6.50

The chapters include: Considering the Audience, Wasteful Prose, Business Jargon, Ideas and Language, Preparation and Organization of Materials, Writing to Persuade, Reading and Listening, Communicating in Group Discussion, etc. The authors have been consultants in communication, and have been engaged for some years in "isolating the problems which cause the businessman to write and speak less effectively than he should", and in identifying the skills and the fundamental principles of communication which the business man needs in order to write and speak well.

Rathbone, Robert R., and James B. Stone / A Writer's Guide for Engineers and Scientists / Prentice Hall, Englewood Cliffs, N.J. / 1962, 348 pp

This book analyzes the problems of a writer, discusses solutions, and provides a variety of models from actual reports. The chapters include: Satisfying the Reader's Needs; Explaining a New Concept, Method, or Device; Developing an Effective

Style; Techniques and Devices to Aid the Writer; etc.

3. Examples of Explanation

Spock, Benjamin / Pocket Book of Baby and Child Care / Pocket Books, New York, N.Y. / 1st printing April 1946, 20th printing June 1951, 502 pp

Full of really excellent examples of explanation. For instance:

> 460. Head Injuries. A fall on the head is a common injury from the age when a baby can roll over (and thereby roll himself off the bed). A parent usually feels guilty the first time this happens. But if a child is so carefully watched that he never has an accident, he is being fussed over too much. His bones may be saved, but his character will be ruined.

Ley, Willy / "The Conquest of Space", with paintings by Chesley Bonestell / The Viking Press, New York, N.Y. / 1949, 160 pp

An excellent example of scientific explanation for an intelligent but lay audience. For instance:

> If the moon's orbit were circular, we would see precisely 50 percent of its surface, some of this area near the rim ill-defined. But the moon's orbit is elliptical, its motion faster when it is nearer earth, and slower farther away. The irregular motion combined with the regular rotation produces a kind of "wobble" which enables us to see temporarily portions "beyond the rim", once on one side, once on the other. Taking advantage of the libration, as the wobble is called in dignified professional language, we can map a total of about four sevenths of the lunar surface. (From page 64)

4. Logic, Fallacies, Symbolic Logic, and Mathematical Systems

Little, W. W., W. H. Wilson, and W. E. Moore / Applied Logic / Houghton Mifflin, Co., Boston, Mass. / 1955, 351 pp

This [illegible] book is the outgrowth of long experimentation with a course in effective thinking which has been for 19 years an integral part of the program of general education at the University of Florida". The book seeks "to develop intellectual skills which the student can put to practical use in solving the myriad problems which will confront him through life." The four parts of the book are: Fallacies; Deduction; Induction; Subjective Factors. The 33 chapters include: Fallacies of Neglected Aspect, The Categorical Syllogism, Truth, Probability, Pitfalls in Thinking, etc.

Black, Max / Critical Thinking: An Introduction to Logic and the Scientific Method / Prentice Hall, Englewood Cliffs, N.J. / 1946, 402 pp

A useful and important book. The three parts are: Deductive Logic; Language; Induction and Scientific Method. The 19 chapters include: The Aims of Logic; Validity and Form; The Uses of Language; Definition; Assorted Fallacies; The Grounds of Belief; Scientific Method. The author says at the start, "I have tried to make this book an argument, not a catalogue of dogmas. Its ideal reader will find himself constantly asking questions, for which he will insist on finding his own answers."

Carnap, Rudolf / Introduction to Symbolic Logic and Its Applications, translated from German into English by William H. Meyer and John Wilkinson / Dover Publications, New York, N.Y. / 1958, 241 pp

A technical handbook and reference manual for symbolic logic, fairly understandable. Part Two is "Application of Symbolic Logic", and contains axiom systems for geometry, physics, space-time topology, and biology.

Allendoerfer, C. B., and C. O. Oakley / Principles of Mathematics / McGraw Hill Book Co., New York, N.Y. / 1955, 448 pp

Chapter 1, "Logic", Chapter 3, "Groups", and Chapter 4, "Sets and Bolean Algebra", give some information and examples of mathematical systems. Most of this book however covers other branches of mathematics, and is outside of the territory of the technique of explanation.

Appendix 5

A Short Glossary

allusion vocabulary — See Class 4 vocabulary.

audience — Persons who are receiving the explanation; the group of explainees.

Class 1, the One-Syllable Vocabulary — Words that are used in an explanation and that are so familiar that every member of the audience will know all of them; for example, "is", "like", "must". Taken to be words of one syllable (unless excluded for some specified reason such as being a special term of the subject); also, words that become two syllables by the addition of one of the endings "-s, -es, -d, -ed, -ing"; also, words for numbers, places, nations, organizations, specific persons, years, dates.

Class 2, the Key Vocabulary — Words that are used in an explanation, that are not specially defined, and that are largely familiar, but perhaps some members of the audience may not have a complete understanding of some of them; for example, "expression", "computer". Taken to be words of two or more syllables, except those classified specifically in Class 3, Class 4, or Class 1.

Class 3, the Special Vocabulary — Words that are used in an explanation, that most members of the audience are not expected to know, and that are (or should be) specially defined and explained in the course of the explanation. The special terms of the subject that occur in the explanation.

Class 4, the Allusion Vocabulary — Words that are used in an explanation, that the audience does not need to know or learn, and that are not used in

later explanations and definitions. Words that occur in an explanation in such a way that understanding them, though helpful, is in no way essential; for example, words occurring in a quotation from a reference when only the general significance of the quotation needs to be understood.

DEC — Digital Equipment Corp. (Maynard, Mass).

explain — To make something clear and plain; to change it from something that is not understood to something that is understood; to render something intelligible; to define or describe a strange idea in terms of familiar ideas and familiar relations.

explainee — Person who is receiving the explanation; a member of the audience.

explainer — Person who is presenting the explanation; lecturer.

explanation — Presenting information clearly and in an interesting way to an audience. Also, the technique or the subject of so doing. Explanation has to be more interesting and to apply more widely than instruction; it is a subdivision of exposition and writing, and it is more general than technical writing, and more general than the presentation of mathematical systems such as geometry. Explanation does not assume a captive audience or an intensely concerned audience.

key vocabulary — See Class 2 vocabulary.

LISP — A certain language for programming computers, which is well adapted to expressing conditions that occur in a great variety of problems, and which is independent of any particular computing machine. LISP is able to deal with letters, digits, numbers, words, statements, commands, lists, and many other items of information composed of symbols of almost any kind. The name LISP comes from the first three letters of "list" and the first letter of "processing". It was worked out in 1958-60 at Massachusetts Institute of Technology, Cambridge, Mass., by John McCarthy and several other computer scientists.

one-syllable vocabulary — See Class 1 vocabulary.

special vocabulary — See Class 3 vocabulary.

understand — In a certain field of knowledge, to acquire (or have) the power to distinguish truth and facts from falsehoods or nonsense, and the power to adapt means to ends, including the power to solve problems.

words for explaining — In an explanation, words that the audience are supposed to know already or to learn while reading the explanation, and that are used as building blocks for later explanation and definition.

Appendix 6

Index

A: accuracy, 47
advice for explaining, short summary of, 179
allusion vocabulary, 61
ambiguous words, 34
assistance by computer, 53
assistance from a computer in explanation, 159
audience, 40
audience for an explanation, 8

B: Basic English, 57
Boolean algebra, 25
brevity, 46

C: captive audience, 14, 17
Churchill, Winston, 50
clarity, demonstration of increase in, 88
Class 1 words, 59
Class 2 words, 59
Class 3 words, 59
Class 4 words, 61
COIK (clear only if known), 36
common properties and relations, 26, 27, 126
communication is subtle: explanation, 245
computer, 147
computer-assisted explanation, 11
computer program for:
    average number of words in a sentence, 160
    detecting key words and phrases, 168
    frequency distribution, 166
    number of syllables, 162
    ratio of number of syllables to number of words, 162
    sentence length, 160
    vocabulary analysis, 67, 181
connotation of words, 37
context, 24
controlling the words for explaining, 63
cybernetics, 148

D: DEC, 67
Digital Equipment Corp., 67
discipline: explanation, 13

E: effectiveness, 51
effort to make good explanation, 156
emotional attitude of explainer, 52
entertainment, 50
exaggeration, removal of, 144
explainee, 8
explanation: definition, 7; distinguishing characteristics, 20; elements, 7;
importance, 10; monetary value, 155
exposition, 16

F: factors affecting explanation, 22
feedback, 51
frequency distribution, computer program for, 166
frequency of words, 37
fuzzy words, 34

G: general plan for improvement of any explanation, 82
giving power to words, 252
grammar, 40

I: importance of an explanation to an audience, 49
improving an explanation, plan for, 82
improving and revising an explanation about:
LISP, 76
statistics, 127
troublemakers, 139
instruction, 14
integral phrases, 62
interest, 50

J: judging and improving an explanation, 75

K: key vocabulary, 60
key words and phrases, computer program for detecting, 168
knowledge of subject matter, 52

L: length of sentence, 43
length of words, 37
Lindbergh, Charles A., 50
LISP: explanation, 76; an original explanation, 93; a revised explanation, 112

M: mathematical systems, 18
mechanical integration, explanation, 209
modifiers, 42
monetary value of good explanation, 155
motivation, 30, 50

N: natural questions to be answered by an explanation, 83
"necessary, required, need", 56
number of words in a sentence, computer program for, 160

O: Ogden, C. K., 57
one-syllable vocabulary, 60

P: paragraphs, 44
PDP-7 computer, 67
personal moral responsibility, 152
personalizing ideas, 30
phoneme, 43
phrases, 39
picture worth a thousand words, 54
plan for improving an explanation, 82
practical aspects of revising explanations, 154
precise truth, 47
precision in long words, 56

R: ratio of number of syllables to number of words, computer program for, 162
readability scale, 161
references of words, 40
resistant sporangia, explanation, 2, 206

S: sampling, explanation of, 66, 217, 224
sentence length, 43
sentence length, computer program for, 160
sentences, 40
ships, explanation of terms, 3, 213
short words, 56
simplicity, 47
sincerity, 52
special terms, 60
special vocabulary, 60
spelling, 32
style, 48
subject: explanation, 13
Swain, Philip W., 252
syllables, rules for number of, 162
symbolic logic, 19

T: taxicab, explanation in a, 1
technical writing, 16
tempo, 47
time available for explaining, 52
trimming: increasing the precision of words and statements, 145
troublemakers, explanation about, 4, 231, 239

U: understanding, 9
utility of the computer program, 88

V: value of computer program for vocabulary analysis, 157
value of good explanation, 154
vividness of words, 38
vocabulary analysis with a computer program, 67, 181

W: well-known words, 55
"we, ours, us", 150
"wimple", 36
words for explaining, 59
wrong words, 33

unclassified
Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Information International Inc. 545 Technology Sq. Cambridge, Mass. 02139	2a. REPORT SECURITY CLASSIFICATION unclassified 2b. GROUP ----

3. REPORT TITLE
Computer-Assisted Explanation; A Guide to Explaining: and Some Ways of Using a Computer to Assist in Clear Explanation

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Published Report, May, 1967

5. AUTHOR(S) (Last name, first name, initial)
Edmund C. Berkeley

6. REPORT DATE May, 1967	7a. TOTAL NO. OF PAGES 280	7b. NO. OF REFS 18
8a. CONTRACT OR GRANT NO. N00014-66-C0281 b. PROJECT NO. --- c. --- d. ---	9a. ORIGINATOR'S REPORT NUMBER(S) 8-217/67 No. 60PB 9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) none	

10. AVAILABILITY/LIMITATION NOTICES
Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES -----	12. SPONSORING MILITARY ACTIVITY Personnel and Training Branch Office of Naval Research, Code 458 Washington, D.C. 20360

13. ABSTRACT

To explain is to make clear or plain, to change something not understood to something that is understood. This document defines explanation; makes clear the differences between explanation and related subjects; describes and illustrates over 50 factors affecting the quality of an explanation; discusses the classes of words in an explanation and gives a program for vocabulary analysis with a computer; explains methods for revising and improving an explanation and gives three illustrations at length; discusses practical aspects of improved explanation; states in full a powerful "Vocabulary Analysis Program" which has been worked out and aids explaining (this program runs on the PDP-7 computer of Digital Equipment Corporation, and a copy of the program tape is available on request); states half a dozen more computer programs for aiding explanation; gives eight sample explanations and their assessment; gives an annotated bibliography of 18 references relating to better explanation; and contains a glossary and an index.

DD FORM 1 JAN 64 1473

unclassified
Security Classification

unclassified

Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
explanation						
explanation techniques						
computer-assisted explanation						
instruction						
computer-assisted instruction						
exposition						
vocabulary analysis						
readability						
computer-assisted vocabulary analysis						

INSTRUCTIONS

1. **ORIGINATING ACTIVITY:** Enter the name and address of the contractor, subcontractor, grantee, Department of Defense activity or other organization *(corporate author)* issuing the report.

2a. **REPORT SECURITY CLASSIFICATION:** Enter the overall security classification of the report. Indicate whether "Restricted Data" is included. Marking is to be in accordance with appropriate security regulations.

2b. **GROUP:** Automatic downgrading is specified in DoD Directive 5200.10 and Armed Forces Industrial Manual. Enter the group number. Also, when applicable, show that optional markings have been used for Group 3 and Group 4 as authorized.

3. **REPORT TITLE:** Enter the complete report title in all capital letters. Titles in all cases should be unclassified. If a meaningful title cannot be selected without classification, show title classification in all capitals in parenthesis immediately following the title.

4. **DESCRIPTIVE NOTES:** If appropriate, enter the type of report, e.g., interim, progress, summary, annual, or final. Give the inclusive dates when a specific reporting period is covered.

5. **AUTHOR(S):** Enter the name(s) of author(s) as shown on or in the report. Enter last name, first name, middle initial. If military, show rank and branch of service. The name of the principal author is an absolute minimum requirement.

6. **REPORT DATE:** Enter the date of the report as day, month, year; or month, year. If more than one date appears on the report, use date of publication.

7a. **TOTAL NUMBER OF PAGES:** The total page count should follow normal pagination procedures, i.e., enter the number of pages containing information.

7b. **NUMBER OF REFERENCES:** Enter the total number of references cited in the report.

8a. **CONTRACT OR GRANT NUMBER:** If appropriate, enter the applicable number of the contract or grant under which the report was written.

8b, 8c, & 8d. **PROJECT NUMBER:** Enter the appropriate military department identification, such as project number, subproject number, system numbers, task number, etc.

9a. **ORIGINATOR'S REPORT NUMBER(S):** Enter the official report number by which the document will be identified and controlled by the originating activity. This number must be unique to this report.

9b. **OTHER REPORT NUMBER(S):** If the report has been assigned any other report numbers *(either by the originator or by the sponsor)*, also enter this number(s).

10. **AVAILABILITY/LIMITATION NOTICES:** Enter any limitations on further dissemination of the report, other than those imposed by security classification, using standard statements such as:

(1) "Qualified requesters may obtain copies of this report from DDC."

(2) "Foreign announcement and dissemination of this report by DDC is not authorized."

(3) "U. S. Government agencies may obtain copies of this report directly from DDC. Other qualified DDC users shall request through _____."

(4) "U. S. military agencies may obtain copies of this report directly from DDC. Other qualified users shall request through _____."

(5) "All distribution of this report is controlled. Qualified DDC users shall request through _____."

If the report has been furnished to the Office of Technical Services, Department of Commerce, for sale to the public, indicate this fact and enter the price, if known.

11. **SUPPLEMENTARY NOTES:** Use for additional explanatory notes.

12. **SPONSORING MILITARY ACTIVITY:** Enter the name of the departmental project office or laboratory sponsoring *(paying for)* the research and development. Include address.

13. **ABSTRACT:** Enter an abstract giving a brief and factual summary of the document indicative of the report, even though it may also appear elsewhere in the body of the technical report. If additional space is required, a continuation sheet shall be attached.

It is highly desirable that the abstract of classified reports be unclassified. Each paragraph of the abstract shall end with an indication of the military security classification of the information in the paragraph, represented as *(TS), (S), (C), or (U)*.

There is no limitation on the length of the abstract. However, the suggested length is from 150 to 225 words.

14. **KEY WORDS:** Key words are technically meaningful terms or short phrases that characterize a report and may be used as index entries for cataloging the report. Key words must be selected so that no security classification is required. Identifiers, such as equipment model designation, trade name, military project code name, geographic location, may be used as key words but will be followed by an indication of technical context. The assignment of links, roles, and weights is optional.